What Is Global Health: Science and Practice Doing to Address Power Imbalances in Publishing?

To begin addressing issues of imbalance and inequity, GHSP is undertaking efforts to identify how we need to do things differently to reflect a range of voices and perspectives in our journal that better corresponds to where global health work is actually being done.

Sonia Abraham, Stephen Hodgins, Abdulmumin Saad, Madeleine Short Fabic

Community Ownership in Primary Health Care—Managing the Intangible

Although enduringly intangible, community ownership is foundational to primary health care. This intangibility is a reminder of what programs can and should do (create space for dialogue, question their own choices, and expand diversity in stakeholder voices making sense of program-induced changes, including through evaluation) and what they cannot do (manage someone else’s ownership).

Eric Sarriot, Ali Nashat Shaar

Counseling Is a Relationship Not Just a Skill: Re-conceptualizing Health Behavior Change Communication by India’s Accredited Social Health Activists

The capacity for India’s community health workers—accredited social health activists—to promote healthy behaviors must be understood within the health system and community context. Their ability to influence health behaviors depends on the strength of their relationships with families and support they receive from the health system.

Rajani Ved, Kerry Scott

Opportunities and Challenges of Delivering Postabortion Care and Postpartum Family Planning During the COVID-19 Pandemic

Devoting scarce health resources to meet the family planning needs of pregnant, postabortion, and postpartum women during the first wave of the COVID-19 pandemic is an investment against higher health systems burdens during subsequent waves of the pandemic and a means to save lives and improve livelihoods.

Anne Pfitzer, Eva Lathrop, Alison Bodenheimer, Saumya Ramachandra, Megan Christofield, Patricia MacDonald, Bethany Arnold, Neeta Bhatnagar, Erin Mielke, Meridith Mikulich
A Qualitative Exploration of Community Ownership of a Maternity Waiting Home Model in Rural Zambia

Community-based maternal child health programs should foster a sense of community ownership to promote sustainability. In rural Zambia, health interventions should be accessible to target communities and clear roles should be established among stakeholders for effective governance.

Constance P. Fontanet, Rachel M. Fong, Jeanette L. Kaiser, Misheck Bwalya, Thandiwe Ngoma, Taryn Vian, Godfrey Biemba, Nancy A. Scott
https://doi.org/10.9745/GHSP-D-20-00136

Bringing Greater Precision to Interactions Between Community Health Workers and Households to Improve Maternal and Newborn Health Outcomes in India

We identified how the quantity and quality of actions taken by community health workers can be refined to move from a one-size-fits-all model to a precision approach that stands to benefit the health of the mothers and newborns they support.

Peter Smittenaar, B.M. Ramesh, Mokshada Jain, James Blanchard, Hannah Kemp, Elisabeth Engl, Shajy Isaac, John Anthony, Ravi Prakash, Vikas Gorthalwal, Vasanthakumar Namasiyayam, Pankaj Kumar, Sema K. Sgaier
https://doi.org/10.9745/GHSP-D-20-00027

Impact of Improved Biomass and Liquid Petroleum Gas Stoves on Birth Outcomes in Rural Nepal: Results of 2 Randomized Trials

Improved biomass stoves may not reduce indoor air pollution as much as is needed to have an impact on adverse birth outcomes.

https://doi.org/10.9745/GHSP-D-20-00011

Are We Using the Right Approach to Change Newborn Care Practices in the Community? Qualitative Evidence From Ethiopia and Northern Nigeria

In Ethiopia, high community-level exposure to consistent messages and the perceptions of community health workers and relationships with them drove newborn care behavior change. In Nigeria, exposure to messages was limited, community health workers were less trusted, and behavior change was reported less frequently.

Zelee Hill, Pauline Scheelbeek, Yashua Hamza, Yared Amare, Joanna Schellenberg
https://doi.org/10.9745/GHSP-D-19-00410
The Critical Role and Evaluation of Community Mobilizers in Polio Eradication in Remote Settings in Africa and Asia

Critical community health worker criteria are important for all community programs, including those focused on a single disease. Areas of importance include community engagement, local adaptation, and linkage with the health system—critical areas for current and future epidemics.

Judy Lewis, Karen LeBan, Roma Solomon, Filimona Bisrat, Samuel Usman, Ahmed Arale
Glob Health Sci Pract. 2020;8(3):396–412
https://doi.org/10.9745/GHSP-D-20-00024

Integrating Calcium Into Antenatal Iron-Folic Acid Supplementation in Ethiopia: Women’s Experiences, Perceptions of Acceptability, and Strategies to Support Calcium Supplement Adherence

In household trials of improved practices, rural Ethiopian women were motivated to adhere to antenatal calcium supplementation regimens, and tailored home-based strategies helped them overcome barriers such as regimen complexity, forgetfulness, side effects, and discouragement from others.

Gina C. Klemm, Zewdie Birhanu, Stephanie E. Ortolano, Yohannes Kebede, Stephanie L. Martin, Girma Mamo, Katherine L. Dickin
Glob Health Sci Pract. 2020;8(3):413–430
https://doi.org/10.9745/GHSP-D-20-00008

What Makes a National Pharmaceutical Track and Trace System Succeed? Lessons From Turkey

Successful implementation of a pharmaceutical track and trace system depended on the political determination to eliminate reimbursement fraud, as well as establishing a pharmaceutical market dominated by a single payer, making reimbursement contingent on verified dispensing and prescription, and being flexible in adapting the system according to stakeholders’ needs.

Koray Parmaksiz, Elizabeth Pisani, Maarten Olivier Kok
https://doi.org/10.9745/GHSP-D-20-00084

Measuring Service Quality and Assessing Its Relationship to Contraceptive Discontinuation: A Prospective Cohort Study in Pakistan and Uganda

The quality of services provided is likely to affect contraceptive continuation. However, findings are strongly influenced by the quality measurement tools used, emphasizing the need for standardization.

https://doi.org/10.9745/GHSP-D-20-00105
Factors Associated With Delayed Contraceptive Implant Removal in Ethiopia

Women receiving implant insertion at the community level were significantly more likely to report keeping their implant for more than 3 years. Even when a referral or back-up system for removals existed, efforts to task-shift the provision of contraceptive implants may have inadvertently led to extended implant use.

Elizabeth Costenbader, Alice F. Cartwright, Misti McDowell, Berhane Assefa, Meza Yirga Tejeji, Eskindir Tenaw
https://doi.org/10.9745/GHSP-D-20-00135

Practical Implications of Policy Guidelines: A GIS Model of the Deployment of Community Health Volunteers in Madagascar

Geographic information systems can be used to support informed decisions about practical issues related to implementing community health worker (CHW) programs. Demands placed on CHWs regarding expected population and surface area coverage and travel time to facilities need to be carefully considered to ensure they are rational and realistic.

Aurélie Brunie, James MacCarthy, Brian Mulligan, Yvette Ribaira, Andry Rabemanantsoa, Louisette Rahantaniarina, Caleb Parker, Emily Keyes
https://doi.org/10.9745/GHSP-D-19-00421

Private Providers’ Experiences Implementing a Package of Interventions to Improve Quality of Care in Kenya: Findings From a Qualitative Evaluation

Although private providers felt that social franchising, quality improvement interventions, and accreditation helped them to increase the quantity and quality of services in their facilities, the quality improvement process was viewed as prohibitively expensive, and the accreditation process often was complex and difficult to navigate without outside assistance.

Masila Syengo, Lauren Suchman
https://doi.org/10.9745/GHSP-D-20-00034

Determinants of Facility-Level Use of Electronic Immunization Registries in Tanzania and Zambia: An Observational Analysis

We provide a framework to quantify the use of electronic immunization registry systems at the facility level and results show the importance of behavioral and organizational factors in explaining their sustained use in Tanzania and Zambia.

Emily Carnahan, Ellen Ferriss, Emily Beylerian, Francis DienMwansa, Ngwegwe Bulula, Dafrossa Lyimo, Anna Kalbarczyk, Alain B. Labrique, Laurie Werner, Jessica C. Shearer
https://doi.org/10.9745/GHSP-D-20-00134
The All Babies Count Initiative: Impact of a Health System Improvement Approach on Neonatal Care and Outcomes in Rwanda

A health system improvement program combining facility readiness support, clinical training/mentoring, and improvement collaboratives increased quality improvement capacity, improved maternal and newborn quality of care, and reduced neonatal mortality. These results can be used to inform system improvement approach design to transform quality of care and outcomes for newborns.

Hema Magge, Evrard Nahimana, Jean ClaudeMugunga, Fulgence Nkikabahizi, Elisabeth Tadiri, Felix Sayinzoga, Anatole Manzi, Merab Nyishime, Francois Biziyaremye, Hari Iyer, Bethany Hedt-Gauthier, Lisa R. Hirschhorn

Glob Health Sci Pract. 2020;8(3):505–517
https://doi.org/10.9745/GHSP-D-20-00031

Where Do Caregivers Take Their Sick Children for Care? An Analysis of Care Seeking and Equity in 24 USAID Priority Countries

Understanding whether and where parents take sick children for care is critical to improve child health and survival. Stakeholders should use this information to ensure that resources are programmed effectively and that sectors complement one another to increase equitable access to high quality integrated management approaches for sick child care.

Sarah E.K. Bradley, Lauren Rosapep, Tess Shiras

https://doi.org/10.9745/GHSP-D-20-00115

A Qualitative Comparative Analysis of the Drivers of HIV Status Knowledge in Orphans and Vulnerable Children in Mozambique

We identified combinations of modifiable factors that HIV programs supporting orphans, vulnerable children, and their families may be able to act on to increase the proportion of beneficiaries who know their HIV status.

Allie Davis, Zola Allen, Nena do Nascimento, Jenifer Chapman, Rotafina Donco, Daan Velthausz

https://doi.org/10.9745/GHSP-D-20-00311

Matching Development of Point-of-Care Diagnostic Tests to the Local Context: A Case Study of Visceral Leishmaniasis in Kenya and Uganda

We provide a new protocol to connect how findings from field research on the local health care setting in resource-limited regions can inform researchers that are working toward developing a new point-of-care diagnostic test for neglected tropical diseases.

Michel Bengtson, Mitasha Bharadwaj, Astrid ten Bosch, Hellen Nyakundi, Damaris Matake-Muhia, Cees Dekker, Jan-Carel Diehl

https://doi.org/10.9745/GHSP-D-20-00028
REVIEWS

Factors That Influence Data Use to Improve Health Service Delivery in Low- and Middle-Income Countries

We identified factors that may influence the relationship between information generation and improvement of health service delivery: governance (leadership, participatory monitoring, regular review of data); production of information (presentation of findings, data quality, qualitative data); and health information system resources (electronic health management information systems, organizational structure, training).

Nicole Rendell, Kamalini Lokuge, Alexander Rosewell, Emma Field
https://doi.org/10.9745/GHSP-D-19-00388

METHODOLOGIES

Mask Reuse in the COVID-19 Pandemic: Creating an Inexpensive and Scalable Ultraviolet System for Filtering Facepiece Respirator Decontamination

We outline a simple, low-cost design—both scalable and adaptable worldwide—to decontaminate filtering facepiece respirators (FFRs) using ultraviolet bulbs and supplies found in most hardware stores. The setup will help health care workers safely reuse FFRs in light of the shortages during the COVID-19 pandemic.

Rachel M. Gilbert, Michael J. Donzanti, Daniel J. Minahan, Jasmine Shirazi, Christine L. Hatem, Brielle Hayward-Piatkovskyi, Allyson M. Dang, Katherine M. Nelson, Kimberly L. Bothi, Jason P. Gleghorn
Glob Health Sci Pract. 2020;8(3):582–595
https://doi.org/10.9745/GHSP-D-20-00218

FIELD ACTION REPORTS

Using Patient-Reported Outcome Measures to Promote Patient-Centered Practice: Building Capacity Among Pediatric Physiotherapists in Rwanda

Tracking outcomes is integral to assessing effectiveness of health systems. Multimodal training was offered in the use of a contextually appropriate, patient-centered outcome measure in a low-resource setting. Results offer insights for designing future capacity-building programs.

Monika Mann, Ines Musabyemariya, Linn Harding, Ben Braxley
Glob Health Sci Pract. 2020;8(3):596–605
https://doi.org/10.9745/GHSP-D-19-00408
RESOURCES

Top 10 Resources in Global Surgery

This resource list could serve to orient those interested in global surgery and could be supplemented with resources advocating for global surgery from clinical, population health, or policy perspectives.

Alliance Niyikuri, Emily R. Smith, Dominique Vervoort, Mark G. Shrim, Stav Brown, Alexander W. Peters, Gavin Yamey, Emmanuel Makasa

https://doi.org/10.9745/GHSP-D-20-00050

CORRECTIONS

Corrigendum: Halperin DT. Coping With COVID-19: Learning From Past Pandemics to Avoid Pitfalls and Panic

Glob Health Sci Pract. 2020;8(3):612
https://doi.org/10.9745/GHSP-D-20-00340
What Is Global Health: Science and Practice Doing to Address Power Imbalances in Publishing?

Sonia Abraham, a Stephen Hodgins, b Abdulmumin Saad, c Madeleine Short Fabric c

What is labeled “global health” has largely concerned the practice of public health work elsewhere, generally in low- and middle-income countries (LMICs). Indeed, global health’s key feature is that its power structures are generally located in high-income countries (HICs) while its implementation is generally located in LMICs. This imbalance is a result of colonial history, funding sources, and social and economic structures that have conferred power—including privilege, prominence, recognition, funding, opportunity, and decision-making authority—to institutions and individuals based in HICs. These deep-rooted structures have helped amplify the voices of those in HICs over the voices of those based in LMICs. In such a system, it is accepted that HICs have expertise to provide and LMICs have capacity gaps to fill. This imbalance is reflected in global health program planning, implementation, research, and publishing. We recognize that they are also reflected at GHSP.

Amplified voice for those based in United States and elsewhere in HICs and diminished voice for those based in LMICs is a poor recipe for improving well-being or strengthening institutions around the world. Indeed, the notion that HICs have something to “teach” LMICs but nothing to learn is a reflection of skewed perceptions of expertise and power. These asymmetries have grown even more evident during the COVID-19 pandemic.

Recent efforts to “decolonize global health” signal an increasing commitment by many players to address these issues of imbalance and inequity. At GHSP, we recognize that to meaningfully engage in addressing power imbalances, as a first step, we must look at our own attitudes and practices. We are especially interested in identifying how we need to do things differently to reflect a range of voices and perspectives in our journal that better corresponds to where this work is actually being done.

1. We completed a critical examination of the demographic composition of our editorial board and associate editorial team. We found that although we have some gender diversity among our editorial team and editorial board, institutions and individuals based in LMICs are underrepresented, as are people of color. We are actively working to ensure diversity in gender, race, ethnicity, and geography of our editorial board, associate editorial team, and peer reviewers.

2. We are committed to ensuring that perspectives from authors based in LMICs are evident in the articles we publish. Any article submitted to GHSP reporting results from research or program experience in specific LMICs should include authors from these countries. In instances where a submission lacks local authorship, under our revised authorship guidelines, the corresponding author is now expected to provide reasons for any such omission in the cover letter accompanying the submission. Additionally, we are working to revise our instructions for authors to better promote ethical authorship practices.

3. We are committed to removing publication barriers that can disproportionately impact authors based in LMICs. Recognizing that journal fees can be a major impediment to article submission and publication, especially for researchers in LMICs, we continue to make GHSP a no-fee, open-access journal. Furthermore, in instances where English language barriers could hinder publication opportunities, our editorial team has worked with authors to address language barriers and is committed to continuing this practice.

4. We expect that any individuals or institutions from HICs that are implementing programs or conducting research in LMICs will respectfully and substantively engage stakeholders from those countries in decision making, planning, implementation, and...
research and are revising our instructions to authors to reflect that expectation.\textsuperscript{12,13}

We recognize that the challenges of power asymmetries and inadequate diversity are complex and multilayered, and our efforts to address them constitute only a modest first step in the needed direction. We commit to holding ourselves to account, and we invite our readers and other stakeholders to hold us to account as we work to meaningfully follow through on these initial actions. To that end, we invite our readers, especially those in LMICs, to share their perspectives through letters to the editor, commentary, and research.

Disclaimer: The views expressed in this article are solely the views of the authors and do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

Competing interests: None declared.

REFERENCES

10. Smith E, Hunt M, Master Z. Authorship ethics in global health research partnerships between researchers from low or middle income countries and high income countries. BMC Med Ethics. 2014;15:42. CrossRef. Medline

Received: September 3, 2020; Accepted: September 3, 2020


© Abraham et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are properly cited. To view a copy of the license, visit http://creativecommons.org/licenses/by/4.0/. When linking to this article, please use the following permanent link: https://doi.org/10.9745/GHSP-D-20-00453
Community Ownership in Primary Health Care—Managing the Intangible

Eric Sarriot, a Ali Nashat Shaar b

Key Messages

- The concept of community ownership in primary health care has a long history but remains challenged in terms of definition, measurement, and differences of perspective from practitioners on a gradient between utilitarianism and empowerment. It continues to be somewhat intangible.
- Although a universal definition across time and contexts may be illusory, contextual appreciation of its dynamic evolution under programmatic influences—for different stakeholders with diverse agendas—is accessible to evaluation and learning.
- No one can “manage” someone else’s ownership, but programs can reject hubris and tokenism by intentionally questioning their unavoidable impact on community ownership and whether they foster it through meaningful dialogue and “sense-making” with local stakeholders.

See related article by Fontanet et al.

In this issue of GHSP, Fontanet et al.1 invite us to return to a concept that has existed since early discussions of community medicine2 and primary health care3: community ownership in health. Many of us who work in global health have felt and seen the excitement and sense of possibility when communities took charge, made a project “their own,” innovated to find contextual solutions, and generated energy and hope in addition to buy-in for a lifesaving or health-promoting intervention. In 1992, one of this article’s authors witnessed how heavy rains had damaged a clinic serving the poor population of Jiftlik in the Jordan valley. Without institutional funds to rehabilitate the structure, the village residents felt a sense of ownership and accountability and restored the clinic themselves, and this clinic is still providing services in 2020. The literature is rich with case studies like this.4–6

As critical as community ownership is—and even foundational for many—it also appears to remain somewhat intangible, possibly impractical for some, and certainly complex for all. We consider some of the reasons for this quandary.

DEFINING COMMUNITY OWNERSHIP

The first stumbling block with community ownership is definitional. This naturally starts with, “what is community really?” This question is followed by—as we generally discuss social processes writ large rather than physical assets7—“what is ownership?” We will satisfy ourselves for now with the idea that a community can be a geographically and demographically defined group of people, a network of people with a common agenda or challenge (illness), and/or most likely a combination of both of these, which creates the possibility of being in a community but outside of important social relationships.

Fontanet et al.1 remind us of the looseness of the concept of community ownership and frame it first under the Paris Declaration of Aid Effectiveness8; community ownership would fit with country ownership, albeit on a different, more local scale. (Oxfam and Save the Children, for their part, see a shift in emphasis from community to country as “a more state-centric form of ownership.”9)

Community ownership is sometimes defined through requirements for ownership, including capacity, empowerment, leadership, value found in the provision of a service, aspirations, and participation, or through consequences of ownership, including participation (again), financial commitment, contributions, and organization membership.10–15 These definitions can sometimes appear tautological—that ownership is defined by the fact of owning or institutionalizing a process or a goal. The literature associates ownership with sustainability of activities and outcomes, a means to achieve cultural adaptation for effective intervention models and to build problem-solving capacity.10,12,16 Ownership can be described as a requirement to build community capacity in a health promotion effort, yet capacity can be presented

---

a Save the Children, Washington DC, USA.
b Palestinian Child Institute, An-Najah National University, Nablus, Palestine.
Correspondence to Eric Sarriot (esarriot@savechildren.org).

Global Health: Science and Practice 2020 | Volume 8 | Number 3
as a requirement of ownership.\textsuperscript{10} Whichever way the causal link is created, it is presented on the path to effective and sustainable health interventions. Countless evaluation reports have also associated failure of achievement and sustainability to the lack of community ownership generated by external projects. In the past, the concept has also been associated to financial contributions by communities,\textsuperscript{17} something critically revised through the universal health coverage agenda.

Much like the concept of participation, ownership lives in the tension between utilitarianism and empowerment,\textsuperscript{18} bridging over to human rights, democratic, and humanist perspectives on development processes. The Ottawa Charter for Health Promotion encouraged a process for enabling communities to increase control over and improve health and notably stated\textsuperscript{19}:

\textit{Health promotion works through concrete and effective community action in setting priorities, making decisions, planning strategies and implementing them to achieve better health. At the heart of this process is the empowerment of communities—their ownership and control of their own endeavors and destinies.}

Advancing community ownership faces at least 3 other challenges.

\section*{Idealistic Framing}

Although we support and believe in the Ottawa Charter’s vision of seeking to increase people’s control over their own health, we must also acknowledge that calls for ownership and “full participation” (as in the recent Astana statement\textsuperscript{20}) sometimes contain an element of idealism that pragmatists can occasionally point out with a wink or with cynicism in the face of harsh “field” realities. Community members may in fact be satisfied sometimes by simply being clients of health services. Demands for social accountability surge when quality, equity, responsiveness, and access conditions are not met. But when they are, people might satisfy themselves with utilizing, rather than owning, a service.

Indeed, public health problems are defined in a context, and these “problems-in-context” demand specific solution configurations, not all of which require the same level of social engagement. People responding to an acute threat might not perceive ownership as an immediate priority. Of course, the global health community had to rapidly re-discover the importance of building a response with communities in the Ebola emergency and efforts to eradicate polio.\textsuperscript{21,22} The current global challenges with vaccine acceptance and the coronavirus disease (COVID-19) situation\textsuperscript{23} are also signaling that some form of ownership is required for scale, sustainability, and impact of interventions. Still, we must also acknowledge that many short-term bets can be won with money and energy invested in proximal determinants of health. Ownership is critical but may be a distal determinant of success. We undermine our own advocacy if we appear to take for granted the value of technicity, policy, and organization in solving health challenges and present ownership in absolute terms.

\section*{Underappreciation for the Inherent Threat to Ownership From External Projects}

Why are we asking about ownership ultimately? Because although they are always well-intended, not infrequently effective, and sometimes sustainable, our external projects inherently displace power and ownership from “natural” social systems (if there is such a thing). We punctuate an equilibrium, if not of ownership, at least of acceptance or resignation to a social baseline, but unless some new equilibrium of ownership is found between diverse stakeholders, the system will be attracted back to its baseline or some other suboptimal state.

Ignoring this tension poses a great risk of hubris. We know the stereotype: experts can come and “give messages,” tell people what the evidence says, and incentivize them to follow their plan, while failing to listen honestly and with respect to the local and community-appropriate ideas for adaptation of the approaches. White elephants are built. Without being a cynic, simply having self-satisfaction with giving token respect for the value of community ownership or coopting can lead to asking the wrong questions, in other words, having a poor definition of what problems really need to be addressed in context. Policy makers close a market to create social distancing; populations protest because they weigh differently an epidemiological risk against the necessity of feeding their family; the market reopens, but no effective community-owned risk reduction solution has been developed.

Although the concern about projects’ displacement of ownership may have been born out of an evolution of international programs away from colonialism, “do-gooding,” and hubris, it also applies to any national or regional program trying to reach remote, poor, minority, or neglected areas. Displacement of ownership is not an
international development problem; it is a universal central-to-local (resource rich to resource poor) development problem. And while “we” question “their” ownership, we are rarely fully accountable for what role and agency we choose to keep to ourselves as we transition.24

MEASUREMENT FOR PROGRAMS
We already mentioned different dimensions through which ownership has been framed. Efforts at measurement naturally must also be multidimensional,9 but this is not the greatest measurement challenge. Research may be able to draw conclusions from a distance on the ownership demonstrated by various communities and stakeholders, but program evaluation—seeking to assess what allows or hinders ownership during implementation—must be carried out with the stakeholders or else be meaningless.

As is the case for assessing institutional capacity, assessing or measuring ownership requires that the “owners” at least acquiesce to the process. A thought experiment can make the point. How would our employers or neighbors react to an outsider knocking on their virtual door to measure their ownership of a stated goal? While accepting to step on the scale does not influence the weight that will be posted on the scale, the measurement of a community’s ownership has community prerequisites in terms of buy-in and boundary decisions (who is the community and who is asking the question?). The prerequisites for measuring ownership are not independent of the ownership variable. It is noteworthy that Fontanet et al. allowed different stakeholders to define their ownership differently. Elements of subjectivity seem unavoidable—not something typically desired in project performance management.

This subjectivity comes with management challenges. Projects try to manage by results and give evidence for achievements. We develop indicators that are as objective and reliable as possible. But when it comes to measuring changes in a social system, our log frames and theories of change are challenged to capture the interaction between our programs and social dynamics over time.25 We say that we “cannot manage it if we cannot measure it,” but given the nature of the question, can we ever manage the ownership of someone else? Then, what are we trying to measure, who should be doing the measurement, and over what timeframe, if ownership evolves on a different timeline than service outputs?

Last and not least, ownership in a complex social system is always changing (dynamic) and can be affected by small changes in interpersonal relationships, services, or operational rules. A new equilibrium between stakeholders comes with new rules and boundaries, and questions may be raised about the ownership allowed for newcomers.11 The stakeholders of community ownership will change, their relationships will change, their perspectives will evolve, as shown by Fontanet et al. over just a 24-month period.

This leaves us with a series of limitations:

- We should assess our impact on community ownership, but our measurement is likely to be subjective and flawed.
- We want to be accountable for progress, but community ownership is precisely about things that we must let go of.
- We should be concerned about community ownership, but we still cannot totally define it. Its local definition depends on who sits around the table. It may change and change substantially based on small evolutions of the problem-in-context.

Should we just abandon all hope? Perhaps not.

CONCLUSION
Social scientists will continue to enrich our understanding by dissecting ownership for different problems and contexts. The measurement challenge may be like that of social capital, for which operational measures can be defined in different contexts, even if a set of universal measures for all contexts may remain out of reach.26 Fontanet et al.1 interestingly circumvent some of the challenges by exploring with qualitative rigor the perceptions of ownership, providing substance to the concept from stakeholders, who have different but compatible definitions of what ownership is to them. The intangible is not made totally tangible, but the local meaning for stakeholders provides guidance to continue developing a program. Another role of research may thus be to provide substance for advocacy and to challenge approaches that deny agency to marginalized communities.

Not all programs have access to strong research capability. However, they can use monitoring, learning, evaluation, and accountability tools to limit disrupting ownership or even to foster it. Promoting community ownership and learning about its development may be more akin to generating new social equilibria than planning for the
delivery of a discrete outcome. It demands genuine interactions, creating enabling conditions and spaces for incremental changes, and building shared values. These ideas are not far from the concept of “harnessing complexity” in complex social and institutional systems. It quite possibly will require monitoring “us”—how we use our money, power, and time, and maybe addressing more critically when we must act and when we must choose to use restraint—as much as measuring “their” ownership. Sustainability-conscious public health practitioners, whether national or international, may not need to worry about precisely measuring the state of community ownership, but to focus more on which agents of the local system are taking agency, how much, and how diverse voices give meaning to tangible changes and intangible perceptions about structures, services, actions, relationships, and values.

If we are intent on finding viable long-term solutions to primary health care challenges with a view of Sustainable Development, transition, and the “journey to self-reliance,” the greatest mistake may be failing to critically engage in questioning our projects’ effects on community ownership and to mistrust the ability of communities to be agents of change.

As messy as it may be.

Acknowledgments: Thank you to Erica Nelson, Lenette Golding, and Judy Sarriot for technical and editorial inputs.

Competing interests: None declared.

REFERENCES


Counseling Is a Relationship Not Just a Skill: Re-conceptualizing Health Behavior Change Communication by India’s Accredited Social Health Activists

Rajani Ved,a Kerry Scottb

Key Messages

- As both community members and health system functionaries, the capacity of India’s community health workers—acredited social health activists (ASHAs)—to improve health outcomes depends on their relationships with families and the support they receive from the health system.
- Training ASHAs on interpersonal communication, such as how to convey the right messages in a persuasive manner, is an important step to improve their impact as health behavior change communicators. However, it is not sufficient.
- Effective counseling also requires that ASHAs have strong positive relationships with community members. These relationships are developed over time as ASHAs link communities to high quality health services and showcase their capacity to directly meet key health needs.

See related article by Smittenaar et al.

Now in its fifteenth year, India’s accredited social health activist (ASHA) community health worker (CHW) program demonstrates all the opportunities and challenges that come with operating CHW programs at a massive scale, in this case with close to a million ASHAs. In this article, we discuss the tensions and benefits associated with ASHAs being both health system actors and community members and how policy and social relationships can support or undermine ASHA ability to improve community health.

The ASHA program is a part of a global re-emergence of national CHW programs,1,2 with the potential of CHWs to contribute to the Sustainable Development Goals around primary health care now well established in the academic literature3 and enshrined in World Health Organization guidelines.4

The article by Smittenaar et al.5 in this issue of GHSP is an important contribution to our understanding of CHW performance and impact for 2 reasons. First, this research examines a mature government CHW program operating at scale rather than of a small pilot or non-governmental organization initiative operating with higher resource inputs in non-generalizable settings. Second, this large-scale survey provides a uniquely rich data set that positions ASHAs within their community environment by assessing indicators from ASHAs, mothers, husbands, and mothers-in-law. The latter 2 categories have not been studied previously at scale in relation to ASHA reach and communication.

As health system actors, ASHAs are only as effective as the system that supports them. Individual-level factors related to ASHA performance, such as ASHA knowledge, must be contextualized within the health system support context, such as the training and continuing supervision they receive. ASHA time-use is closely tied to the health system context because ASHAs will decide whether to perform a role based on financial incentives, transportation considerations, competing demands, and also their self-assessment of their capacity to meet the family’s needs and expectations.

The article reported that receiving home-based newborn care was associated with positive health behaviors (clean cord care and exclusive breastfeeding). The study also found that only 31% of mothers were receiving adequate home visits in the first week, and ASHAs spent only 8.8% of their time on home visits to postpartum women. This care provision gap is striking, considering that home visits are 1 of the 5 key roles identified for ASHAs, alongside coordinating Village Health and Nutrition Days, convening the Village Health and Sanitation Committees, accompanying patients to health facilities, and maintaining basic records. The 2013 ASHA guidelines include clear directions for when ASHAs should visit postpartum women. Financial incentives for providing home-based newborn care are quite substantial and are being

---

a National Health Systems Resource Centre, New Delhi, Delhi National Capital Region, India.
b Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA.
Correspondence to Kerry Scott (kscott26@jhu.edu).
claimed by a significant proportion of ASHAs; for example, in FY 2019–2020, data from the state’s reporting system show that about 76% of Uttar Pradesh’s ASHAs reported carrying out home-based newborn care visits. Further investigation is needed to understand the contradiction between this study’s finding of low coverage and the policy-level emphasis on ASHAs conducting home-based newborn care. Low coverage was also identified as a challenge in a study on postnatal home visits that analyzed large-scale surveys. The key questions to our mind relate to health systems functionality. Are ASHAs skipping home visits due to a lack of equipment or training that has undermined their motivation and makes them feel that they have little to offer during these visits? If an ASHA lacks a weighing scale, thermometer, and watch, or is not confident that she can adequately assess a newborn, she may feel there is no point in performing a home visit. To what extent does this contradiction call for bolstered accountability and oversight of ASHAs? Although the financial incentive that ASHAs receive for performing home-based newborn care is an important policy lever, it is clearly insufficient in ensuring that these visits occur.

As community members, ASHA relationships and interactions with families are subject to the broader power relations that shape social norms and behaviors. There is great value in studying CHWs as social actors engaged within power systems, as the Smittenaar et al. manuscript does. Just as ASHAs work within a health system context, they also work within community systems where they must navigate social relationships mediated by gender, age, caste, and other social hierarchies. Smittenaar et al. highlight some of the ways in which ASHAs navigate these relational dynamics, especially around targeting counseling to husbands and mothers-in-law. The broader community relationships and norms that ASHAs navigate require further exploration. Caste dynamics play out across almost all facets of rural life and may also influence ASHA acceptance into family homes or willingness to visit certain families. Community norms that prohibit outside visitors from interacting with newborns may be a major barrier facing ASHAs when trying to carry out home-based newborn care visits.

Many ASHAs have accrued strong social capital in their communities, and are seen as a vital link to the health system. An ASHA’s ability to maintain this social capital hinges in part on the health system’s performance, such as how families are treated by health facility staff during childbirth after they have been encouraged by the ASHA to have an institutional birth. ASHA efforts to maintain relationships and build social capital may be manifested in the large amount of time spent on accompanying women to health facilities. When the ASHA program began in 2005, ASHAs needed to physically attend a birth to receive their Janani Suraksha Yojana incentive payment. However, this policy changed some years later. ASHAs do not need to accompany women in labor; instead women just have to confirm that their ASHAs supported them in having an institutional delivery. And yet, as shown in the analysis by Smittenaar et al., the practice of accompanying women in labor persists and is associated with higher quality of care for women. Thus, this ASHA behavior endures beyond financial motivation and appears grounded in ASHA efforts to secure good care for women, maintain relationships, and meet community expectations.

As India moves toward primary health care reform, ASHAs are being mobilized to expand beyond maternal, child, and reproductive health into noncommunicable disease and mental health care. Although financial incentives for this additional work will be required, they must be coupled with the training on interpersonal communication called for by Smittenaar et al. But it is important that interpersonal communication be understood within the health system support and community context. Effective counseling is not just about messages or communication styles. It is not just a skill that can be taught. Instead, effective counseling occurs within a relationship of trust between the ASHA and beneficiary, which is built over time as ASHAs link up to functional health services and accrue social capital in the community. This is particularly true for entrenched normative behaviors, such as newborn care or diet or those related to stigmatized conditions, such as around mental health. ASHAs are best able to improve community health when they are skilled counselors who also earn community trust through directly providing some health services and supporting beneficiaries in accessing good quality higher-level care.

Competing interests: None declared.

REFERENCES


Received: August 31, 2020; Accepted: August 31, 2020

Cite this article as: Counseling is a relationship not just a skill: re-conceptualizing health behavior change communication by India’s accredited social health activists. *Glob Health Sci Pract*. 2020;8(3):332-334. https://doi.org/10.9745/GHSP-D-20-00426

© Ved and Scott. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are properly cited. To view a copy of the license, visit http://creativecommons.org/licenses/by/4.0/. When linking to this article, please use the following permanent link: https://doi.org/10.9745/GHSP-D-20-00426
Opportunities and Challenges of Delivering Postabortion Care and Postpartum Family Planning During the COVID-19 Pandemic

Anne Pfitzer,a Eva Lathrop,b Alison Bodenheimer,c Saumya RamaRao,d Megan Christofield,a Patricia MacDonald,e Bethany Arnold,a Neeta Bhatnagar,a Erin Mielke,e Meridith Mikuliche

Key Messages

We seek to assist decision makers in maximizing provision of essential services without compromising access to quality family planning care and while minimizing the risk of COVID-19 transmission among clients, and between clients and health care workers.

- Managers should help facility teams to integrate counseling and provide a range of contraceptive methods as is feasible within existing contacts with pregnant, postabortion, birthing, and postpartum women, even as services migrate to new models with a mixture of in-person and virtual/tele-health consultations.
- Policy makers should prioritize devoting resources to meet the family planning needs of pregnant, postabortion, birthing, and postpartum women, and the health care workers serving them as an investment against higher health systems burdens in later months and during subsequent waves of the pandemic.

Le tableau et encadre de dialogue de l’article est aussi disponible en français.

INTRODUCTION

The coronavirus disease (COVID-19) pandemic is poised to cause infection and death in millions of people across the globe at a stunning pace. The scale of the required response will inevitably pivot attention and resources toward fighting the pandemic and away from essential reproductive, maternal, newborn, and child health care, including access to voluntary family planning in the critical postabortion and postpartum periods. Decreased access to these lifesaving services will lead to a downstream increase in maternal and child morbidity and mortality. Data from previous complex emergencies demonstrate that a decrease in access to family planning results in increased poor outcomes related to unintended pregnancies and abortions. In a pandemic as vast and unique as COVID-19, where the primary mitigating factor is elimination of close physical contact, harnessing opportunities to provide family planning education, services, and supplies while women are already interfacing with the health care system during pregnancy and the postabortion, childbirth, and postpartum periods is strategic and lifesaving. This will require acceleration of integrated service delivery as well as creative and dynamic innovations of alternative service delivery approaches to address the family planning needs of pregnant, postabortion, birthing, and postpartum women. Investment in documentation of programmatic learnings could offer insights and opportunities for improving the resilience of health systems. Devoting scarce health resources to ensure the family planning needs of pregnant, postabortion, birthing, and postpartum women are met during the first wave of the COVID-19 pandemic is an investment against higher health systems burdens in later months and during subsequent waves of the pandemic and a means to ultimately save lives and improve livelihoods.

WHY FOCUS ON FAMILY PLANNING WITHIN SERVICES FOR PREGNANT, POSTABORTION, AND POSTPARTUM WOMEN NOW?

Closely spaced and unintended pregnancies are a public health concern and can have detrimental effects on women, infants, and children. The World Health Organization (WHO) recommends a 24-month interval from live birth until subsequent pregnancy to reduce the risk of adverse maternal, perinatal, and infant outcomes. Similarly, a woman who experiences a miscarriage or induced abortion that requires emergency treatment will rapidly...
return to fertility, with ovulation within 14–28 days depending on gestation.11 Thus, offering voluntary family planning counseling and services as part of postabortion care (PAC) as well as during pregnancy, childbirth, and the postpartum period is a critical means to protect vulnerable postpartum and postabortion women and reduce unintended and closely spaced pregnancies.12,13 Both postpartum and postabortion provision of voluntary contraceptive counseling and services constitute high-impact practices.14 In this commentary, we refer to the broad care of postabortion women as PAC, which includes15:

- two essential services: (1) treatment of emergency complications, and (2) voluntary family planning counseling, including provision of contraception.

We use postabortion family planning (PAFP) when specifically referencing that component of PAC.

Before the COVID-19 pandemic, unmet need for modern contraception among women who wished to delay, space, or limit future childbearing and were not currently using a method in the postpartum and postabortion periods was already high.16–18 Because the COVID-19 pandemic has affected both supply- and demand-side access to family planning, women’s ability to achieve their reproductive intentions has been further compromised. Health systems worldwide seek to reduce facility visits to protect the health workforce and clients from the spread of COVID-19. Individuals’ health-seeking behavior is changing too, as they avoid facilities or seek care from alternate sources because of fear of acquiring the infection, respect for distancing measures, and/or mobility restrictions.

It is precisely here where postpartum family planning (PPFP) and PAFP offer a unique opportunity to make the most of facility and pharmacy visits and interactions with community health workers that individuals continue to have during the COVID-19 crisis. Now more than ever, and as others have already pointed out,19 the care that pregnant, postpartum, and postabortion women receive could be optimized to also meet their family planning needs by integrating contraceptive counseling and services for those who wish to space or limit their next pregnancy and to yield the significant health and well-being aims of voluntary contraceptive uptake and healthy timing and spacing of pregnancies.

Thus, country health programs and facilities must prepare now and for the future to serve pregnant, postabortion, birthing, and postpartum women’s needs and to ensure women and their accompanying partners are informed, educated, and counseled on voluntary contraception. Although currently available guidelines center on infection prevention and control and immunization services, several international bodies have clarified the essential nature of family planning and maternal health care including the WHO (see these resources20 and these21).

## PRACTICAL APPROACHES AND MODIFIED FAMILY PLANNING SERVICE DELIVERY MODELS FOR DIVERSE POSTABORTION AND POSTPARTUM CLIENTS

To facilitate a reduction in COVID-19 transmission without compromising the quality of voluntary PAFP and PPFP counseling and services, we must enhance safe delivery of existing integrated service models (drawing lessons from previous emergencies) and also implement innovative, alternative service delivery mechanisms.

The key pillars in WHO’s strategy to reduce human-to-human transmission that must be incorporated into these models include:

1. Social distancing (e.g., supporting stay-at-home orders, limiting crowds in facilities, and reducing number of patient-provider contacts)
2. Early identification and isolation of cases (e.g., systematic screening, contact tracing, and community-based referral pathways)
3. Infection prevention and control (e.g., hand hygiene, appropriate personal protective equipment, and cleaning supplies)

We encourage countries to follow the WHO operational considerations22 for case management of COVID-19 and share WHO or locally adapted risk communication materials23 across all health sectors.

Health systems must confront rapidly changing challenges to maintain provision of essential health services, including PAC and PPFP. Overcoming these challenges will require a high level of intersectoral collaboration, communication, transparency, and community engagement. Specific obstacles will vary in number and magnitude by setting and may be particularly burdensome in already fragile settings. Early data highlight the high personal toll on health care workers (HCWs) during the pandemic in terms of their own physical health and risk of contracting the virus, mental health, structural concerns (availability/lack of personal protective equipment, long hours, etc.).24 Policy makers and program managers must identify which disruptions most significantly affect
family planning outcomes so that limited resources can be allocated most effectively. These obstacles may include, but are not limited to:

- Supply chain disruptions for family planning and infection prevention and control products (e.g., stock-outs)
- Clinic closures, reduced hours, and diminished capacity to treat high client volumes
- Redeployment of skilled family planning health staff to COVID-19 response
- HCW fear of contagion/contamination and attacks by community members
- Client hesitancy to access public transportation; health facilities treating COVID-19 patients and other crowded facilities (e.g., pharmacies, waiting rooms)
- Mobility or movement restrictions impacting clients and some cadres of the health workforce
- Income loss among individuals and families to access or pay user fees for contraceptive services
- Adoption of recommendations for limiting attendance of support people at in-person antenatal care (ANC), delivery, or postnatal visits; shifting of selected ANC visits to telehealth; and early postpartum discharge from facilities.

On the last point, the recommendations inhibit or delay attributes of care in normal times, such as joint decision making around PPFP and may reduce time for PPFP counseling and method provision. Similar restrictions in PAC inhibit or delay male engagement in counseling and joint family planning decision making.

We highlight several approaches to address these challenges and maximize opportunities for voluntary PPFP and PAFP counseling and services (Table). Optimal voluntary adoption of PPFP and PAFP will be achieved not only through integration during provision of routine maternal and newborn care, but also by integrating family planning into other essential service contacts and outreach mechanisms.

These recommendations, of course, must be tailored to each unique setting for both logistical and cultural purposes. We also recognize that HCWs, especially in fragile settings, face myriad challenges during normal times which are only exacerbated during this pandemic. We hope HCWs recognize that maximizing opportunities with a client reduces the need for return visits and consequently the risks to themselves and their peers. Also, it enhances care for their clients in that it reduces the need to expose themselves to additional risks associated with separate family planning visits. Adjustments are required not only within health facilities, but throughout the health systems.

We hope that program managers use the recommendations (Table) to promote these efforts in their communications with HCWs. This will encourage both HCWs and clients to feel safer, regardless of where they are seeking family planning information, products, and services. Thus, a blend of facility-based, community-based, and virtual/telehealth services could be used per setting, as context, health system, and community capacity allow. Additionally, providing clients with timely and accurate anticipatory guidance regarding changes to routine health care services will be essential in supporting their continued access to family planning. We encourage systems to monitor trends in utilization of various services along the continuum of care, at multiple levels, from facility to district to regional to national. We also encourage program managers to recognize HCWs and health facility teams who problem solve and innovate to optimize integration of services and suggest they document and disseminate process improvements and modifications so as to encourage appropriate replication.

### LEARNING RELATED TO PAC AND PPFP DURING COVID-19

Providers, policy makers, and those in positions of leadership can rely to some extent on past experiences in complex emergencies such as Ebola, Zika, and humanitarian responses to guide practice and service delivery in the context of the COVID-19 pandemic, but there remains much we do not know. Unique features of this disease can influence care differently than outbreaks of the past. For example, little is known about the impact of COVID-19 on pregnancy and postpartum recovery or the ways women’s and families’ health-seeking behaviors may change in the face of this pandemic. All provider cadres will likely experience a tension between duty of care and self-protection, and what educational messages, training, and protection strategies will work for them remains unknown. Others have highlighted that shifts toward self-care or short-acting contraceptive methods may have ripple effects, hopefully temporary, on the global supply chain for contraception that merit close monitoring. It is an imperative of the response community to 

---

**Policy makers and program managers must identify which disruptions most significantly affect family planning outcomes so that limited resources can be allocated most effectively.**

---

Maximizing opportunities with a client reduces the need for return visits and consequently the risks to clients, HCWs, and their peers.
**TABLE. Recommendations Related to Integrating Postabortion and Postpartum Family Planning for Diverse Categories of Women**

<table>
<thead>
<tr>
<th>Population</th>
<th>What PPFP or PAFP Service Is Relevant Here?</th>
<th>How Might This Service Be Affected by COVID-19 Mitigation?</th>
<th>How Might You Modify the PPFP or PAFP Services in Context of COVID-19?</th>
<th>Key Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women seeking ANC services</td>
<td>PPFP counseling at every ANC contact (whether in person or virtual) and messages reinforced at each subsequent contact to help women formulate a plan for voluntary PPFP.</td>
<td>ANC schedules and visits may be modified to allow for screening, triage, scheduling, social distancing, shifting some visits to virtual platforms or teleconsultations, providing ANC through CHWs, and/or by combining ANC contacts.</td>
<td>All Women: PPFP counseling at each of these ANC contacts remains paramount, particularly as disruptions to ANC may result in inconsistent models of care and providers throughout the pregnancy and beyond. Staff and clients follow local guidelines for wearing masks, especially when social distancing cannot be observed.</td>
<td>COVID-19 Technical Brief for Maternity Services (UNFPA)26</td>
</tr>
<tr>
<td>Women presenting for PAC</td>
<td>PPFP counseling to include all contraceptive methods and information on return to fertility. Voluntary FP service provision of all eligible methods.</td>
<td>Due to supply and/or service limitations, possible shifts to teleconsultations for medical management of abortion complications; the ability to provide a full range of FP options during PAC could be limited.</td>
<td>All Women: Where provision of a woman’s desired contraceptive method is not immediately possible, counsel about alternatives and plan for obtaining their preferred method once services and supply stabilize. When providing methods, observe all infection prevention and control protocols. For long-acting reversible methods, ensure client has a plan for managing side effects and for obtaining removal services. For short-acting methods, provide advance prescriptions and refills for several months depending on stock availability.</td>
<td>WHO recommendations on antenatal care for a positive pregnancy experience (includes PPFP ANC counseling)27</td>
</tr>
<tr>
<td>Women seeking maternity services</td>
<td>PPFP counseling to include all contraceptive methods and information on return to fertility, relative to exclusive breastfeeding practices. Voluntary service provision of all eligible methods.</td>
<td>After uncomplicated vaginal delivery, and where the home setting is suitable for recovery, health systems may consider early discharge with frequent telehealth monitoring for routine postnatal care and referrals for postpartum or newborn emergency care.</td>
<td>All Women: In these instances, continue offering PPFP counseling prior to discharge, emphasizing return to fertility and how and where to access PPFP during later postnatal or immunization visits. When providing FP methods, observe all infection prevention and control protocols. For long-acting reversible methods, ensure client has a plan for managing side effects and for obtaining removal services. For short-acting methods, consider advance dispensing ECPs and/or condoms for LAM users, or POPs or progesterone vaginal rings for those who will breastfeed, but not exclusively.</td>
<td>COVID-19 Technical Brief for Maternity Services (UNFPA)26</td>
</tr>
</tbody>
</table>

**Continued**
<table>
<thead>
<tr>
<th>Population</th>
<th>What PPFP or PAFP Service Is Relevant Here?</th>
<th>How Might This Service Be Affected by COVID-19 Mitigation?</th>
<th>How Might You Modify the PPFP or PAFP Services in Context of COVID-19?</th>
<th>Key Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Woman presenting for postnatal care and infant immunization services in facilities</td>
<td>PPFP counseling at PNC and immunization contacts. Voluntary FP service provision of all eligible methods.</td>
<td>Return postnatal care visits had low coverage in normal times and may be further compromised by COVID-19. Immunization coverage, while typically higher in normal times, is also affected by COVID-19.</td>
<td>All Women: Strengthen counseling during return postnatal care (PNC), and where appropriate, within immunization services. If privacy does not allow FP counseling during immunization/well baby clinic and staffing adequate, provide intra-facility linkage for FP for the mother. Consider utilizing lay counselors where clinical staff are overburdened or there are no staff providing FP services concurrently to immunization. In group education sessions, keep clients seated 2 meters apart; staff and clients follow local guidelines for wearing masks. <strong>Additional for women with positive COVID-19 test or symptoms:</strong> Same as with ANC, maintain adequate separation from asymptomatic women. Provide or reiterate mother with COVID-19 specific guidance for safe breastfeeding. Plan for additional COVID-19 related support.</td>
<td>Guiding principles for immunization activities during the COVID-19 [32] Family Planning and Immunization Integration [33]</td>
</tr>
<tr>
<td>Pregnant, delivering and postpartum women not accessing facility-based services</td>
<td>PPFP and PAFP counseling at every community-based contact (whether in person or virtual). Information on return to fertility, support for exclusive breastfeeding practices where applicable. Voluntary FP service provision and/or referral of all eligible methods.</td>
<td>Community- and home-based services may be utilized in greater magnitude due to limitations on facility-based care.</td>
<td>All Women: Emphasize PPFP and PAFP counseling and information on how to access services as part of CHW-led and other community-based initiatives, observing privacy and confidentiality. Health education platforms can also enhance education on PPFP and PAFP and educate women on how to use fertility awareness (with caution for women who have yet to see regular menstrual cycles return) and self-administered methods, such as LAM, POPs, and where available, progesterone vaginal rings. Various community stakeholders (women’s groups, community leaders, faith community) can also encourage essential services including PAC and PPFP, and provide support for finding these services.</td>
<td>Risk Communication and Community Engagement (RCCE) Action Plan Guidance COVID-19 Preparedness and Response [34] Community-based health care, including outreach and campaigns, in the context of the COVID-19 pandemic [35]</td>
</tr>
</tbody>
</table>
explore gaps in our knowledge on both the health system and user sides, develop research protocols to generate answers, and document learning to inform ongoing care as COVID-19 continues to be a part of the new global reality (Box).

**TABLE.** Continued

<table>
<thead>
<tr>
<th>a General principles for all women:</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Promote respectful, stigma-free care, with cautious communication to blame the virus, not the person. This applies to women experiencing abortion complications as well.</td>
</tr>
<tr>
<td>- Promote task-sharing where CHWs already exist, including to encourage referrals to facility services and build trust through risk communication and community engagement.</td>
</tr>
<tr>
<td>- In the immediate postpartum or postabortion periods, women are, by default, nonusers after pregnancy. After method adoption, women may also need support to manage side effects and/or removal. Refer to contraception and COVID-19 guidance in such cases.</td>
</tr>
<tr>
<td>- Be on the lookout for signs of gender-based violence and support victims, as incidence of violence is expected to increase.</td>
</tr>
<tr>
<td>- Post or adapt WHO infographics on gender-based violence for displaying in health facilities and consider providing training to health workers who work with women on the 2016 WHO ANC recommendation which suggest clinical inquiry and referrals for GBV.</td>
</tr>
</tbody>
</table>

**b Factors that may influence implementation of PAC and PPFP services during the COVID-19 pandemic:**

- COVID-19 testing availability
- Country-specific laws, clinical guidelines and practice standards
- Task-sharing practices
- Pervasiveness of mobile phones and other communication technologies to facilitate “telehealth”
- Availability of IPC supplies (hand hygiene resources, personal protective equipment including masks for all HCWs and clients), environmental cleaning, and waste management
- Stay-at-home orders and/or curfews
- Modifications to ANC, PAC, childbirth, PNC and immunization services
- Religious or cultural practices

Abbreviations: ANC, antenatal care; CHWs, community health workers; ECPs, emergency contraceptive pills; EmONC, emergency obstetric and newborn care; FP, family planning; GBV, gender-based violence; HCW, health care workers; IAM, Lactational Amenorrhea Method; LARC, long-acting reversible contraception; PAC, postabortion care; PAFP, postabortion family planning; PPFP, postpartum family planning; PNC, postnatal care; POPs, progestin-only pills; Q&A, question and answer; RH, reproductive health; UNFPA, United Nations Population Fund; WHO, World Health Organization.

**LINKAGES BETWEEN RESPONSE ON PAC AND PPFP AND HEALTH SYSTEM RESILIENCE**

The emergence of COVID-19 has tested health systems worldwide, both in their management and mitigation of the pandemic directly, but also in their ability to maintain essential services for their populations. The WHO notes in the COVID-19 Operational Guidance for Maintaining Essential Health Services they:

> a system’s ability to maintain delivery of essential health services will depend on its baseline capacity and burden of disease

alongside their COVID-19 transmission context. Thus, it is health systems’ resilience—or their capacity to prepare for and effectively respond to crises, maintain core functions when a crisis hits, and adapt and transform to function effectively post-pandemic—that offers a route to stymie COVID-19’s deleterious effects on essential health services both now and in subsequent waves of the pandemic.

It is impossible to ignore the threats of not taking action. Based on experience from previous epidemics and health system shocks, we recognize that both family planning and maternal, newborn, and child health (MNCH) care and outcomes also stand to lose ground. One analysis of maternal and reproductive health outcomes estimates that a 10% decline in the use of essential care will result in 1.7 million additional women and 2.6 million additional newborns who will experience major complications as a direct result of care disruptions. Further, a 10% decline in modern contraceptive use would result in nearly 50 million additional women with unmet need for contraception. Amidst the Ebola outbreak in West Africa in 2014, maternal health stakeholders saw their coverage of ANC, facility delivery, and PNC drop. As health seeking patterns amidst COVID-19 appear to echo those witnessed during Ebola, experts estimate a similar, yet more substantial loss now—one which results in hundreds of thousands of additional child and maternal deaths.
Amidst this gloomy outlook, MNCH programs are managing to provide services because pregnant women still need them. The ability to deliver these services comes in part due to rapid adaptations to provide safe care at community and household levels (including through self-care). Now, more than ever, the clarion call for integration of family planning with essential MNCH care appears: in the context of limited health service accessibility, optimizing every contact to uphold the health, well-being, and interests of women for their health and the health of their families. Simply put, PPFP and PAC integrate services to respond to individuals’ multidimensional needs with an array of simultaneous health interventions (in this case, voluntary family planning linked with maternal and/or infant health care). The health system adaptations we seek now and as we look to the future are both reactive to the moment we live in and an investment in the resilience of the system for the future. Opportunities for integration are central to—and should be capitalized upon—even in the midst of a crisis.

CONCLUSION

The ability of women, girls, and couples to freely choose the number, timing, and spacing of their pregnancies is a fundamental right and a means to achieve multiple sustainable development goals. Global actors have called for family planning to remain on the list of essential services during the COVID-19 pandemic, along with other key maternal, newborn, and child health care services. PAC and postpartum family planning intersect multiple categories of essential services. Prioritizing integrated service provision now promises to reap returns for improved health and well-being by preventing a rise in closely spaced pregnancies that may require care and burden facilities during subsequent waves of the epidemic. In the months to come, we can cultivate health system resilience by incorporating innovative models of integrated service delivery for pregnant, postabortion, delivering, and postpartum women; securing resources for programs to innovate and sustain services; and seeking partnerships between communities and MNCH programs and across the public and private sectors.

Acknowledgments: The impetus for this article came from a discussion of the FP2020 PPFP/PAFP Steering Committee, which provides strategic guidance to the global postpartum and postabortion family planning movement building on momentum from the 2015 Global PPFP Meeting in Chiang Mai, and continues under the management of the FP2020 Secretariat, with the ultimate goal of supporting progress at the country level.
Disclaimer/Funding statement: The contents are the sole responsibility of the authors and do not necessarily reflect the views of the United Nations Population Fund (UNFPA) or the U.S. Government.

Competing interests: None declared.

REFERENCES


46. Townsend JW, ten Hoope-Bender P, Sheffield J. In the response to COVID-19, we can’t forget health system commitments to contraception and family planning. Int J Gynaecol Obstet. Published online May 16, 2020. CrossRef. Medline

Peer Reviewed

Received: June 12, 2020; Accepted: August 18, 2020; First published online: September 17, 2020


© Pfitzer et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are properly cited. To view a copy of the license, visit http://creativecommons.org/licenses/by/4.0/. When linking to this article, please use the following permanent link: https://doi.org/10.9745/GHSP-D-20-00263
A Qualitative Exploration of Community Ownership of a Maternity Waiting Home Model in Rural Zambia

Constance P. Fontanet, Rachel M. Fong, Jeanette L. Kaiser, Misheck Bwalya, Thandiwe Ngoma, Taryn Vian, Godfrey Biemba, Nancy A. Scott

Key Findings
- Community members’ perceptions of ownership were related to their ability to use the maternity waiting home (MWH) and feeling a sense of responsibility toward its success.
- Representing their community’s interests was a crucial component of the role of the governance committee and management unit.
- Collaboration between the governance committee and the health facility staff was key to allowing the MWH to meet its goal of serving the community.

Key Implication
- To improve the sustainability of community-based maternal and child health programs, program managers should ensure that interventions are accessible to target communities and clear roles are established among stakeholders.

ABSTRACT

Context: Ownership is an important construct of sustainability for community-based health programming, though it is often not clearly defined or measured. We implemented and evaluated a community-driven maternity waiting home (MWH) model in rural Zambia. We engaged stakeholders at all levels and provided intensive mentorship to an MWH governance committee comprised of community-selected members. We then examined how different stakeholders perceive community ownership of the MWH.

Methods: We conducted 42 focus group discussions with community stakeholders (pregnant women, fathers, elders, and community health volunteers) and 161 in-depth interviews with MWH stakeholders (health facility staff, district health officials, and MWH governance committee and management unit members) at multiple time-points over 24 months. We conducted a content analysis and triangulated findings to understand community ownership of the MWH and observe changes in perceptions of ownership over time.

Results: Community members’ perceptions of ownership were related to their ability to use the MWH and a responsibility toward its success. Community and MWH stakeholders described increasingly more specific responsibilities over time. Governance committee and management unit members perceived their ability to represent the community as a crucial component of their role. Multiple respondent types saw collaboration between the governance committee and the health facility staff as key to allowing the MWH to meet its goal of serving the community.

Conclusion: The perceptions of community ownership evolved as the intervention became more established. Use of the MWH, and clear understanding of roles and responsibilities in management of the MWH, seemed to foster feelings of community ownership. To improve the sustainability of community-based maternal and child health programs, interventions should be accessible to target communities and clear roles should be established among stakeholders.

INTRODUCTION

To address the underlying causes of maternal mortality and morbidity, governments throughout sub-Saharan Africa (SSA) have implemented both supply-side and demand-side interventions. Many interventions build capacity of community members through community health worker training or use community volunteers to conduct health education and promotion activities. In Zambia, community...
members have been engaged in the implementation of maternity waiting homes (MWHs), which are residential dwellings located near health facilities where women can stay to await childbirth by a skilled birth attendant and receive postnatal care services. Women who can access a health facility with a high-quality MWH are more likely to deliver at a facility where a skilled birth attendant is present. Additionally, this type of intervention has shown promising results on reducing mortality among pregnant mothers in Africa. However, in several studies of MWHs, women and community members were concerned about the sustainability of the intervention, regardless of its perceived success.

Evaluating the sustainability of externally funded interventions to address maternal mortality is increasingly important in global health. A recent review found that very few health interventions in SSA examined sustainability outcomes. Of those that did, the majority identified community ownership and mobilization as crucial facilitators of intervention sustainability. For example, an evaluation of the large-scale, comprehensive Saving Mothers Giving Life initiative, which aimed to rapidly reduce maternal mortality in Zambia and Uganda from 2012 to 2016 through community health worker mobilization, doctor and nurse training, and facility upgrades, found that the intervention’s ability to foster a sense of local ownership around the intervention was an essential factor in the early maintenance of its gains in maternal and child health outcomes.

The 2005 Paris Declaration on Aid Effectiveness and the 2008 Accra Agenda for Action both argue in favor of a country defining its own development priorities and designing and leading programs promoting these priorities. The need for country or community ownership of health interventions, for example through sustained government or ministry of health funding or community contributions, is rooted in ideological values of self-determination and has been posited to be an effective approach to sustainability. The concept of community ownership emerged in the literature several decades ago from similar ideological origins. In the health context, community ownership has been defined as “community leaders’ levels of perceived control over key functions of a [health program] at the time of measurement.” Ultimately, empowerment is the desired outcome, where communities have control and decision-making ability over these health interventions and their future, whether at the local (community ownership) or national level (country ownership).

When conducting formative research to design a community-driven MWH model in rural Zambia, we found that community members considered community ownership essential to the success of a potential MWH intervention. More specifically, community members linked the concept of sustainability of the MWH intervention to local or community ownership but did not offer clear definitions or examples of what ownership meant to them. In the evaluation of the sustainability of our MWH model in rural Zambia, we used qualitative methods to explore local perceptions of community ownership over the course of the MWH intervention. This article qualitatively explores how different stakeholders perceived community ownership of the MWH and how this changed over the first 24 months of MWH operations.

### METHODS

#### Study Setting/Intervention Design

The Maternity Homes Access in Zambia project constructed 10 MWHs adjacent to rural health centers able to provide obstetric care for uncomplicated deliveries and within 2 hours of time to a referral hospital equipped to care for women experiencing obstetric complications. The intervention was implemented in 4 districts of rural Zambia: Choma, Pemba, and Kalomo (in Southern Province) and Nyimba (in Eastern Province). All study districts are primarily rural with some peri-urban pockets. Choma has 247,860 people, 76% of whom live in rural areas. At the time of the 2010 census, Pemba was part of Choma. Kalomo has 258,570 people, most of whom live in a rural area (93%). Nyimba has 77,359 people, 91% of whom live in rural areas.

We gathered community input from community members and relevant stakeholders in the health system and traditional leadership structures to design an intervention that would meet community standards of acceptability. The resulting 3-pillar conceptual model (core MWH model) focused on: (1) the establishment of quality MWH structures with functional infrastructure and amenities; (2) the need for a community-based system to oversee the daily management, finances, and future maintenance requirements of the MWHs without overburdening the existing health system; and (3) the need to be linked with the health system for clinical care of waiting women and education. The core MWH model met
cultural-appropriateness and was aligned with Ministry of Health policy.8,9,23

In accordance with the management pillar, we engaged stakeholders at all levels of the MWH ecosystem before and during the intervention implementation (Figure). We engaged community members, including traditional leadership (i.e., chiefs and the village headmen who represent the chiefs), to sensitize them on the benefits of an MWH, actively participate in the governance and management of the MWH through selected community members, and contribute to the financial and operational sustainability of the MWH. We also engaged health system staff, which included staff at the health facility and district health office levels, to ensure our goals were aligned. For example, we engaged district health staff to participate in steering committees to advise the creation of the MWH governance committees and MWH management units. We engaged health facility staff and community health outreach workers to actively participate in the governance and management of the MWHs and to ensure linkage of the MWH to the facility.

We provided training and ongoing mentorship to community-elected MWH governance committees and management units (GCMU). The governance committees are comprised of community members and health facility staff. The management units are comprised of community members or health facility staff selected by the governance committees. The governance committee is responsible for managing the MWH, mobilizing resources, and overseeing the management unit to ensure sustainability of the MWH. The management unit is responsible for the daily operations of the MWH and management of MWH assets. Additionally, we covered the start-up costs for community-led income-generating activities that could help support costs associated with the MWH and contribute to its financial sustainability. The project phased out supporting implementation in April 2018 but continued to monitor intervention activities through October 2018. The evaluation of the implementation of the intervention has been described elsewhere.25

**Thematic Framework**

To evaluate the sustainability of our MWH intervention, we relied on findings from the formative evaluation and Scheirer and Dearing’s framework for the sustainability of public health programs.26 The framework determines sustainability by asking: (1) whether program activities were continued after external support ends, (2) whether community-level partnerships or coalitions developed during the funded program were maintained, and (3) whether new organizational practices, procedures, and policies that were started during program implementation were maintained. We hypothesized that community ownership may be an important mediator of these constructs and therefore an essential component of sustainability. This hypothesis was in line with findings from our formative work, which qualitatively underscored the importance of ownership of the MWH intervention by the community.4,8,9 We deliberately did not define community ownership, but rather allowed our stakeholders to explain ownership in their own words.
Data Collection and Management
We conducted 42 focus group discussions (FGDs) and 161 in-depth interviews (IDIs). The FGDs were conducted with 412 community members (14 groups of pregnant or recently delivered women, 10 groups of men with a child under age 1, 9 groups of community elders, and 9 groups of community health volunteers). Safe Motherhood Action Groups made up the majority of community health volunteers, but traditional birth attendants were also part of the FGDs. The IDIs were conducted with MWH governance committee and management unit members, and health systems staff (health facility staff, district health officials). FGDs were conducted at 3 timepoints: immediately following intervention launch (October 2016 to January 2017); during the intervention (August 2017 to September 2017); and after implementation phaseout (April 2018 to May 2018). IDIs were conducted at 4 timepoints: immediately following intervention launch (October 2016 to January 2017); during the intervention (April 2017 to June 2017 and November 2017 to January 2018); and after implementation phaseout (July 2018 to October 2018). We used convenience sampling to select the most senior person available on the day of visit for the district staff, health facility staff, governance committee, and management unit IDIs. Community health volunteers recruited FGDs participants from varying distances from the health facility. Both qualitative instruments captured basic demographics and had questions that elicited perceptions of the MWH operations and stakeholder roles as well as perspectives on health facility engagement, community ownership, and long-term sustainability.

Local data collectors fluent in English and the local languages, who were trained in qualitative interviewing techniques, the interview guides, and research ethics, administered the IDIs and FGDs. Data collectors were not members of the intervention implementation team, which provided direct mentorship and support to the GCMU, as described above. Data collectors participated in a refresher training before each round of qualitative interviews. Predefined probes were adapted and refined based on results from each previous round. IDIs and FGDs were audio recorded, translated into English, and transcribed verbatim into Microsoft Word.

Analysis
Transcripts were systematically coded in NVivo version 11 (QSR International). The main coding nodes were identified a priori based on the questions and probes in interview guides. Transcripts were double coded against the theoretical framework and to a topic or theme. Additional nodes were added as themes emerged during coding. We conducted a content analysis to assess respondent definitions of community ownership and applicability to the MWH intervention among respondent types and over time.

Demographic data were captured in SurveyCTO Collect version 2.212 (Dobility, Inc.) and analyzed in SAS version 9.4 (SAS Institute Inc.). Proportions were calculated for respondent sex, occupation, and school attendance. Means and standard deviations (SD) were calculated for respondent age and highest grade completed. We had missing data (n=24) for years of education for elders and community health volunteers at project phaseout.

Ethics
We obtained ethical approval through the Boston University Medical Campus Institutional Review Board and the ERES Converge Institutional Review Board in Lusaka, Zambia, and approval by the Zambian National Health Research Authority. Written informed consent was obtained from respondents in the language they were most comfortable using: English, Chinyanja, or Chitonga.

RESULTS
We have provided a description of IDI respondents (Table 1) and FGD respondents (Table 2). Results are presented by the 2 main themes that respondents discussed: (1) general perceptions of ownership of the MWH, and (2) roles and responsibilities for each stakeholder toward the MWH.

We conducted 42 FGDs with 412 individuals over 24 months (Table 1). FGD respondents were fairly similar across time points. Most had attended at least some schooling although community elders had less than other respondents. Community health outreach workers and community elders were more likely to be male. Pregnant and recently delivered women were slightly younger and had 3–4 live births, and men had 4–5 children.

Within the health system and MWHs, 161 IDIs were conducted (Table 2). The majority of MWH staff (management unit and governance committee) interviewed were female and the majority of health system staff respondents were male. The health system staff had been in their current...
positions a few years, working in the broader health system for much longer.

**Theme 1: General Perceptions of Ownership**

Generally, respondents agreed at all time points that the community had an important ownership role in the MWH. However, community members (FGD respondents) described ownership from the point of view of potential or real users and MWH stakeholders (IDI respondents), who are part of the health system, described ownership in terms of roles and management.

FGD respondents perceived that they—the community—owned the MWH and described 2 different elements of ownership: (1) the ability for any member of the community to use the MWH, and (2) a sense of responsibility for the future success of the MWH. A sample of illustrative quotes are included in Supplement 1. Across all time points, respondents justified that the MWH belonged to the community because any member could use it during their pregnancy. Specifically, pregnant women talked about being able to stay in the MWH and use its amenities, therefore being the owners of the MWH:

| TABLE 1. Demographic Characteristics of Focus Group Discussion Respondents (N = 412) on Ownership of Maternity Waiting Homes in 4 Districts in Zambia |
|-------------------------------------------------|-------------------------------------------------|-------------------------------------------------|-------------------------------------------------|
|                                | Immediately Post-launch | Implementation Period | Project Phaseout |
|                                | October 2016 to January 2017 | August 2017 to September 2017 | April 2018 to May 2018 | Total |
| Pregnant/recently delivered women | n=46 | n=34 | n=40 | n=120 |
| Age, y, mean (SD)              | 26 (7) | 25 (6) | 25 (6) | 25 (6) |
| Pregnant, No. (%)              | 20 (43.5) | 19 (55.9) | 22 (55.0) | 61 (50.8) |
| Education, y, mean (SD)        | 7 (3) | 7 (3) | – | 7 (3) |
| Parity, mean (SD)              | 3 (2) | 3 (2) | 3 (2) | 3 (2) |
| Gravida, mean (SD)             | 3 (2) | 3 (2) | 4 (2) | 3 (2) |
| Married/cohabiting, No. (%)    | 37 (80.4) | 28 (82.4) | 32 (80.0) | 97 (80.8) |
| Men with child under 1 year old | n=46 | n=36 | n=16 | n=98 |
| Age, y, mean (SD)              | 33 (12) | 34 (9) | 30 (9) | 33 (11) |
| Education, y, mean (SD)        | 8 (3) | 9 (3) | 5 (4) | 8 (3) |
| Number of biological children, mean (SD) | 5 (3) | 5 (3) | 4 (2) | 4 (3) |
| Married/cohabiting, No. (%)    | 46 (100.0) | 35 (97.2) | 14 (87.5) | 95 (96.9) |
| Elders                          | n=46 | n=38 | n=16 | n=100 |
| Female, No. (%)                | 29 (63.0) | 17 (44.7) | 9 (56.3) | 55 (55.0) |
| Age, y, mean (SD)              | 79 (20) | 63 (9) | 64 (9) | 70 (17) |
| Years of education, y, mean (SD) | 5 (4) | 6 (4) | – | 5 (4) |
| Number of biological children, mean (SD) | 7 (3) | 8 (4) | 7 (3) | 7 (3) |
| Married/cohabiting, No. (%)    | 34 (73.9) | 27 (71.1) | 10 (62.5) | 71 (71.0) |
| Community health volunteersa   | n=46 | n=40 | n=8 | n=94 |
| Female, No. (%)                | 23 (50.0) | 29 (72.5) | 5 (62.5) | 57 (60.6) |
| Age, y, mean (SD)              | 53 (19) | 44 (10) | 46 (11) | 49 (16) |
| Education, y, mean (SD)        | 9 (2) | 9 (2) | – | 9 (2) |
| Number of biological children, mean (SD) | 7 (2) | 5 (3) | 4 (3) | 6 (3) |
| Married/cohabiting, No. (%)    | 35 (76.1) | 25 (62.5) | 6 (75.0) | 66 (70.2) |

Abbreviation: SD, standard deviation.

a Safe Motherhood Action Group.
The MWH is for every person, but to be specific, the owners are the pregnant women because they are the ones that use it. —Pregnant woman, Project phaseout

Many stated that not having cost associated with usage made them feel like they owned it. Others reported that not experiencing discrimination fostered a sense of ownership.

Although all community members emphasized the importance of being able to use the MWH, in particular, men, community health volunteers, and elders mentioned a dimension of responsibility for the MWH as they described ownership. Specifically, men and community health volunteers described needing to look after or take care of the MWH:

It is ours in the sense that the users are the community, so it’s the community’s responsibility to take care of it. If anything gets damaged it’s the community to take care of it. —Man, Implementation period

Elders described that having made a financial contribution as a community to the initial construction of the MWH bolstered their sense of ownership because money came from the community to maintain the MWH.

### TABLE 2. Demographic Characteristics of In-depth Interview Respondents (N=161) on Ownership of Maternity Waiting Homes in 4 Districts in Zambia

<table>
<thead>
<tr>
<th>Management unit (MWH staff)</th>
<th>Immediately Post-launch October 2016 to January 2017</th>
<th>Implementation Period April 2017 to June 2017</th>
<th>Project Phaseout November 2017 to January 2018</th>
<th>Project Phaseout July 2018 to October 2018</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>n=8</td>
<td>n=10</td>
<td>n=10</td>
<td>n=9</td>
<td>n=37</td>
<td></td>
</tr>
<tr>
<td>Female, No. (%)</td>
<td>8 (100.0)</td>
<td>8 (80.0)</td>
<td>6 (60.0)</td>
<td>9 (100.0)</td>
<td>31 (83.8)</td>
</tr>
<tr>
<td>Age, y, mean (SD)</td>
<td>41 (1.5)</td>
<td>39 (14)</td>
<td>39 (12)</td>
<td>37 (10)</td>
<td>39 (13)</td>
</tr>
<tr>
<td>Education, y, mean (SD)</td>
<td>10 (2)</td>
<td>10 (2)</td>
<td>10 (2)</td>
<td>10 (2)</td>
<td>10 (2)</td>
</tr>
<tr>
<td>Farmers, No. (%)</td>
<td>4 (50.0)</td>
<td>9 (90.0)</td>
<td>4 (40.0)</td>
<td>8 (88.9)</td>
<td>25 (67.6)</td>
</tr>
<tr>
<td>Governance committee (MWH staff)</td>
<td>n=17</td>
<td>n=18</td>
<td>n=16</td>
<td>n=10</td>
<td>n=61</td>
</tr>
<tr>
<td>Female, No. (%)</td>
<td>11 (64.7)</td>
<td>9 (50.0)</td>
<td>10 (62.5)</td>
<td>5 (50.0)</td>
<td>35 (57.4)</td>
</tr>
<tr>
<td>Age, y, mean (SD)</td>
<td>50 (5)</td>
<td>47 (8)</td>
<td>49 (7)</td>
<td>49 (13)</td>
<td>49 (8)</td>
</tr>
<tr>
<td>Education, y, mean (SD)</td>
<td>9 (2)</td>
<td>10 (2)</td>
<td>9 (2)</td>
<td>10 (2)</td>
<td>9 (2)</td>
</tr>
<tr>
<td>Leadership position in governing committee, No. (%)</td>
<td>5 (29.4)</td>
<td>10 (55.6)</td>
<td>10 (62.5)</td>
<td>9 (90.0)</td>
<td>34 (55.7)</td>
</tr>
<tr>
<td>Farmers, No. (%)</td>
<td>16 (94.1)</td>
<td>17 (94.4)</td>
<td>13 (81.3)</td>
<td>8 (80.0)</td>
<td>54 (88.5)</td>
</tr>
<tr>
<td>Health facility staff</td>
<td>n=11</td>
<td>n=10</td>
<td>n=10</td>
<td>n=10</td>
<td>n=41</td>
</tr>
<tr>
<td>Female, No. (%)</td>
<td>5 (45.5)</td>
<td>6 (60.0)</td>
<td>3 (30.0)</td>
<td>2 (20.0)</td>
<td>16 (39.0)</td>
</tr>
<tr>
<td>Facility in-charge, No. (%)</td>
<td>7 (63.6)</td>
<td>4 (40.0)</td>
<td>2 (20.0)</td>
<td>6 (60.0)</td>
<td>19 (46.3)</td>
</tr>
<tr>
<td>Clinical position, No. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical officer</td>
<td>1 (9.1)</td>
<td>1 (10.0)</td>
<td>2 (20.0)</td>
<td>1 (10.0)</td>
<td>5 (12.2)</td>
</tr>
<tr>
<td>Nurse/midwife</td>
<td>6 (54.6)</td>
<td>7 (70.0)</td>
<td>2 (20.0)</td>
<td>2 (20.0)</td>
<td>17 (41.5)</td>
</tr>
<tr>
<td>Non-skilled birth attendant staff</td>
<td>1 (9.1)</td>
<td>1 (10.0)</td>
<td>3 (30)</td>
<td>2 (20.0)</td>
<td>7 (17.1)</td>
</tr>
<tr>
<td>Years working in the health system, mean (SD)</td>
<td>14 (10)</td>
<td>10 (7)</td>
<td>6 (8)</td>
<td>10 (8)</td>
<td>10 (8)</td>
</tr>
<tr>
<td>District health officers</td>
<td>n=6</td>
<td>n=9</td>
<td>n=3</td>
<td>n=4</td>
<td>n=22</td>
</tr>
<tr>
<td>Female, No. (%)</td>
<td>2 (33.3)</td>
<td>3 (33.3)</td>
<td>1 (33.3)</td>
<td>1 (25.0)</td>
<td>7 (31.8)</td>
</tr>
<tr>
<td>District Health Officer, No. (%)</td>
<td>3 (50.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>3 (13.6)</td>
</tr>
<tr>
<td>Years working in the health system, mean (SD)</td>
<td>11 (7)</td>
<td>11 (5)</td>
<td>9 (6)</td>
<td>16 (4)</td>
<td>12 (6)</td>
</tr>
</tbody>
</table>

Abbreviations: MWH, maternity waiting home; SD, standard deviation.
Because we suffered to build it, we are supposed to take care of it. Because everyone took part in the building of the MWH, everyone feels it belongs to them. So taking care of it, everyone is ready to do that. —Elder, Immediately post-launch

In the IDIs, GCMU respondents consistently described the MWH as belonging to the community and/or pregnant women. Over time, GCMU respondents also better articulated what specifically was being “owned” and better described their own roles and responsibilities in relation to the MWH. Respondents increasingly described material assets and IGA revenue as belonging to the community, but explained that it was earmarked for the care of the MWH or pregnant women and managed by the governance committee.

District staff discussed the MWH as belonging to both the community and the health facility from the outset:

[The MWH] is for the whole community in conjunction with the [health] facility. The greater part of the ownership is [shared] by the community because they’ve been involved in construction, even bringing materials, and when they were being launched. The community was involved, so they know that this is our structure, because it’s built for us. —District health officer, Immediately post-launch

Health facility staff’s perceptions evolved over time. Although health facility staff discussed shared ownership of the MWH with the community, during later rounds of interviews, they more clearly articulated the role of the health facility, culminating with the MWH being described as an extension of the health facility:

The MWH is part of the clinic. When it came to electrification, we were using the same meter. We are even using the same water. —Health facility staff member, Project phaseout

Theme 2: Stakeholder Roles and Responsibilities
FGD and IDI respondents described the specific MWH-related roles for each stakeholder with increasing specificity over time (Table 3). Respondents generally agreed on the role of the health facility staff, but some nuances emerged for the roles of the community at large and the district health staff. Illustrative quotes are summarized in Supplement 2.

Role of the Community
Traditional leadership played an increasingly important role in the functioning of the MWH within the community over time. At the launch of the program, community health volunteers mentioned that the role of the village headmen was to regularly check on the MWH on behalf of the chiefs. Other respondent types such as the governance committee, health facility staff, and community members described traditional leadership as responsible for mobilizing community contributions for the MWH (described below) based on requests from the governance committee (Table 3).

The GCMU members highlighted that the community was the primary owner of the MWH and described that the community at large was responsible for making cash and/or in-kind contributions to the MWH to support maintenance. Community members felt they had a responsibility to make cash and/or in-kind contributions to the MWH:

[Those of us] who come here, we do the cleaning on our own … just the way we do it back at our home. —Pregnant woman, Immediately post-launch

Health facility staff and the GCMU felt the primary role for pregnant women was to use the MWH and sometimes assist with cleaning tasks. They also felt the community at-large should participate in the structural maintenance of the MWH.

Community members only vaguely described community contributions at the launch of the intervention but became increasingly specific over time. Pregnant women described having some responsibility toward helping keep the MWH clean. Respondents from the community described how community contributions of cash, food, and building materials ensured sustainability. However, many pregnant women said that community contributions were not always happening as planned, whether monetary or in kind, and expressed concern over the future of the MWH.

Finally, the GCMU and health facility staff described that the community had a mandate to ensure the future of the MWH. For example, health facility staff described that the community had control over the selection of governance committee members, which were in turn responsible for the success of the MWH. Therefore, the community owns the MWH but delegates its role to the governance committee:

The community is there to ensure that their structure is community-driven. They should ensure, because the governance committee is chosen by the community, so the community will oversee whether those people are...
doing what is expected. They are monitors to ensure that the MWH is there. If a member of the governance committee is not working well, we’ll call the community and plan on how they can help the member or the committee or they can change the committee. The community has the mandate to change the committee because this is a structure that is going to benefit the community. These mothers that are coming here are coming from the villages and the community. The community is going to enjoy and should not shun coming to deliver from the clinic because of the conditions that are not good here.—Health facility staff member, Project phaseout

Role of the Governance Committee and Management Unit

Over time, all respondent types described managerial and custodial roles for both the governance committee and management unit with increasing specificity (Table 3). At the launch of the intervention, most respondent types failed to describe clear structures for the committees. Later, governance committee members, health facility staff, and community members described the governance committee as responsible for managing the income of the MWH and for communicating with the village headmen about the MWH’s needs and possible community contributions. The management unit members described themselves as responsible for the day-to-day operations of the MWH:

| TABLE 3. Focus Group Discussion Respondents’ Perspectives on Stakeholders’ Maternity Waiting Home Roles and Responsibilities Over Time in 4 Districts in Zambia, October 2016 to October 2018 |
|-----------------------------------------------|--|-----------------------------------------------|
| **Immediately Post-launch** | **Implementation Period** | **Project Phaseout** |
| Community at-large | Community members | ● Need to contribute money | ● Pregnant women help with cleaning of MWH | ● Pregnant women help with cleaning of MWH and contribute to IGAs |
| MWH Management Staff | Governance committee | ● Representative of communities | ● Representative of communities | ● Representative of communities |
| | | ● Partner with health facility | ● Responsible for MWH management | ● Responsible for MWH management |
| | Management unit | ● Representative of communities | ● Help maintain the cleanliness and comfort of the MWH | ● Take care of day-to-day MWH needs |
| Health Systems Staff | District staff | ● Respond to health facility needs, but not MWH needs | ● Respond to health facility needs, but not MWH needs | ● Respond to health facility needs, which sometimes include MWH needs |
| Health facility staff | | ● Provide cleaning supplies | ● Partner with the GCMMU for MWH management | ● Partner with the GCMMU for MWH management |
| | | ● Work with GC | ● Provide clinical care | ● Some participate in GC |
| | | ● Provide clinical care | ● Check in on mothers at the MWH | ● Provide clinical care |
| | | ● Check in on mothers at the MWH | ● Communicate with the district | ● Check in on mothers at the MWH |

Abbreviations: GC, governing committee; GCMMU, governing committee/management unit; MWH, maternity waiting home.
be vandalized, and it will not last long. —Management unit staff member, Project phaseout

Community members highlighted the role of the management unit as responsible for explaining the rules of the MWH to expecting mothers and ensuring they felt welcome upon arrival.

Governance committee members described a hierarchical relationship between themselves and the management unit. The management unit is responsible for escalating MWH issues to the governance committee when needed; however, no exact mechanism was described by GCMU members. Overall, all respondent types described the GCMU members as representatives of the community who work as partners to manage the MWH. Additionally, GCMU respondents perceived themselves and the committee structures as important representatives of the community.

We are the ones who are supposed to see to it that all is working accordingly because this MWH belongs to us. If there is anything happening, we communicate with the rest of the community to inform them. —Governance committee member, Implementation Period

Health facility staff referred to governance committee members as custodians of the MWH, including its assets and IGA-generated income, on behalf of the community during later rounds of interviews. In essence, the community “owns” the MWH and the governance committee, which is made up of community-selected members to represent the interests of the community with respect to the MWH operations. Ultimately, governance committee members considered themselves responsible to the community for achieving the mission and goals of the MWH by providing pregnant women with a high-quality MWH, but indicated a reliance on the management unit and the community at-large.

Role of the Health Facility Staff
During the implementation of the project, all respondent types described health facility staff as having a clinical role in caring for the pregnant women utilizing the MWH. They regularly visit the pregnant women to check on their health and monitor potential pregnancy-related complications.

The responsibilities we have mainly concern the mothers. There are some who have overstayed, so we go through the antenatal bookings they have attended [and if we find they were] not fully examined, we will go through that to see if our findings are okay. We refer them together with the management to a higher-level hospital. Apart from that, we still encourage them if there are any questions or problems. We ask them to still come because it is, we are still one facility. —Health facility staff member, Implementation period

Many health facility staff echoed this description and one gave the example of having midwives checking on expecting women at the MWH while the GCMU ensured the MWH remained clean.

Over time, the role of the health facility staff evolved into a supporting role for the GCMU (Table 3). At the launch of the intervention, the health facility staff expressed that they felt responsible for providing the MWH with cleaning supplies because the MWH was not yet generating income. Gradually, and as the income-generating activities were implemented, the management and operations of the MWH were assumed by the GCMU, in collaboration with the health facility. Community members, including recently delivered women, provided a similar description of the evolution of the health facility staff’s role over time. At the launch of the intervention, community members described health facility staff such as nurses, as responsible for the management of the MWH. Over time, respondents began to describe the GCMU and the health facility staff as partners in the operations of the MWH meant for the community. The district staff described that the health facility staff were also responsible for communicating with the district about the MWH’s needs.

Role of the District Staff
Over time, the district staff consistently described themselves as playing a supervisory role based on health facility requests (Table 3). District staff highlighted their need to be responsive to financial and operational issues of the health facility and the MWH but offered no specifics. For example, a district staff member explained that the district staff could assist with providing resources to help the MWH operations continue, but did not specify what might warrant their assistance.

I think one of our major roles here at the district is to help the facility and the community to handle some of the major problems that they face that they are not able to handle at their level and maybe to source support from others or maybe provide resources which can help them run. There are issues which they can handle on
Community Ownership of a Maternity Waiting Home Model in Zambia

Discussion

This study qualitatively explored community ownership of MWHs from the perspective of multiple stakeholders over 24 months, from launch of the intervention to after external support for the program had ended. A core MWH model was designed in consultation with local stakeholders, community leaders, and community members throughout Zambia.4,8,9,23,25 During these consultations, participants acknowledged the need for communities to contribute to the operations and maintenance of the MWH and stated that community involvement or “community ownership” was crucial to MWH sustainability, which has been corroborated by studies in other areas.9

Community ownership is a challenging concept to define because different stakeholders bring varying perspectives and the term is often conflated with other concepts related to community engagement or sustainability.32 When asked about who owned the MWH, respondents all agreed that the MWH belonged to the community, but differed in how and what they described as “ownership.” Although respondents in the community described ownership in terms of the MWH being available to them and used by community members, respondents within the health system linked ownership to responsibility, similarly to what has been reported in the literature by implementers of other community-driven interventions.28

Community members focused on describing how they felt the MWH belonged to them because the MWH was built for them. Specifically, community members emphasized the importance of everyone being able to use the MWH, especially because it was free of charge. This finding is interesting considering the country-wide decision to eliminate user fees for maternal health services in Zambia, even if existing evidence does not clearly link the absence of fees with increased utilization.29–31 Our results indicate that not having to pay a fee to use the MWH may have influenced community members’ decision to participate in the intervention and was an important determinant of whether community members felt a sense of ownership over the MWH.

Although the first finding gives insight into what may be a necessary component to foster a sense of community ownership, our second finding was that respondents connected the concepts of community ownership and sustainability with stakeholder roles and responsibilities. Rather than describing their perception of the MWH from the point of view of a user or potential user, respondents who were involved in the operations and management of the MWH focused on how the role of each stakeholder was essential to the overall functioning of the MWH and its future sustainability. Over time, respondents increasingly described more specific roles for other stakeholders.

The role of the community became more precise over time. At first, the community was expected to contribute, but the nature of the contributions and the mechanism through which contributions could be made remained vague. At later timepoints, respondents described that the community was responsible for making cash and/or in-kind contributions. These contributions were to be mobilized from community members by traditional leadership based on feedback from the governance committee.

The district staff were responsible for supervising the health facility staff, but the health facility staff, governance committee, and management unit were essential to the functioning of the MWH. Specifically, the health facility provided clinical care to pregnant women, supported the GCMU, and communicated with government-level actors such as the district staff. The GCMU had different roles in the operations of the MWH, with the governance committee playing a managerial role in the MWH and the management unit being responsible for the day-to-day activities of the MWH. Members of both groups were considered representatives of the community. As such, the concepts of community ownership and clear roles for the stakeholders involved in the management of the MWH were linked together by the respondents.

While respondents initially discussed ownership of the MWH as falling to either the community or the community and the health system, when probed further, specific roles and responsibilities for stakeholders at multiple levels were identified by all respondents as critical to the long-term success of the MWH. The description of these roles with increasing specificity over time is interesting, as respondents were able to articulate responsibilities and things that are needed for the MWH to function now and in the future as they gained more experience managing and using the MWH.
This increasing specificity is to be expected as the MWH operated for nearly 2 years at the time of the last data collection point. All roles became clearer to the stakeholders involved, not only their own roles but also the roles of the other stakeholders. These specific roles and the emphasis on their importance persisted even when external financial and mentorship support from the project staff ended.

Other factors such as social context, financial support, and organizational partnerships affect the sustainability of the intervention, but the perspectives of our respondents are important because they illustrate how the concept of community ownership can be operationalized to serve the purpose of sustainability. In summary, users described ownership as an ability to stay at the MWH and benefit from it, whereas stakeholders involved in the operations of the MWH described ownership as a well-established set of mechanisms where each stakeholder had a specific role widely known to other stakeholders.

Respondents within the community and within the health system described that the MWH intervention created an ecosystem of shared responsibility around the current and future functioning of the MWH (Figure). The GCMU’s roles were to manage the operations of the MWH in collaboration with the health facility, and that they had a responsibility toward the community to ensure that the MWH was providing quality services and that its operation was sustainable. The collaboration between the GCMU and health facility staff represents a critical element of the intervention, the point at which the health facility staff accepted to assume a level of responsibility for the MWH and its functioning, including through communication with the district. This shared responsibility between the community-driven GCMU and the health facility staff reduces the managerial burden of the staff and allows for focus on their other duties. In previous findings, health system stakeholders had similarly described the importance of the GCMU because its presence allows health system staff to attend to their clinical duties. Although the community representatives on the GCMU are the instrumental link to the community, the health facility staff is the instrumental link to the district and government-level actors. This governing body where community representatives and health facility staff formally collaborate to ensure the smooth operations of the MWH serves as a key opportunity to connect the community with the formal health system.

Existing literature indicates that social accountability is an important factor of community participation, especially to ensure equity and a high level of quality in service provision. This type of accountability allows for the successful collaboration of stakeholders, which is the case here. One caveat of this concept is that if community participation does not yield positive results, community members may become less engaged. In our setting, respondents unfortunately did not describe clear mechanisms of accountability between the governance committee and the community, or the governance committee and the management unit, or the GCMU and the health facility. While respondents gave some examples of how these stakeholders communicate, they did not give many examples of successful conflict resolution. In that regard, the lack of described mechanisms for feedback is worrisome. Surprisingly, respondents did not see the annual meetings held by the GCMU with community members as a mechanism for raising and resolving issues. However, this may be the result of not probing for examples of conflicts during the IDIs and FGDs or because interviews and FGDs took place too soon after external support had ended for serious concerns to have emerged. It is also possible that the respondents interviewed in our sample were not among the annual meeting participants. Consensus around roles and responsibilities of members is essential for governance models like the GCMU or health facility committees. Our results add to the evidence showing a link between specified roles and effectiveness of a health intervention. However, our data do not allow us to fully understand how these roles will continue to be maintained, especially when conflicts inevitably arise. This represents an area of both improvement and exploration, as other studies have elicited the difficulties of understanding what works and does not work after external support is removed in interventions that involve communities and community representation at the facility level.

During the design of this intervention, community respondents had stated that community ownership was essential to the sustainability of the intervention, but they did not provide clear parameters for ownership. What emerges from our findings is that community representation through a structure like the governance committee may be an adequate response to the community’s demand for community ownership. The governance committee is concerned with representing the interests of community members, who themselves identified ease of access to the
MWH as an essential component of ownership. Through the core MWH model, the governance committee and the health facility staff became partners in the operations of the MWH, with each party able to facilitate communication with the community and the district about MWH needs. The health facility staff was a link to the formal health system, whereas the GCMU served as a link to the broader community surrounding the MWH.

Participation of the most marginalized members of the community in these representative processes is a concern. Many interventions have sought to improve the effectiveness and sustainability of health programs by engaging community members, especially those directly affected by the health programs. Our intervention was able to achieve broad engagement with the community at large, including our target population of pregnant/recently delivered women. Unfortunately, our findings do not provide information on whether the poorest, most vulnerable pregnant women within the MWH catchment areas felt the same amount of ownership as those who might be considered less vulnerable. Further studies should continue to explore which strategies work best to ensure equity within processes that aim to increase community participation and ownership perhaps by focusing on who is chosen to represent the voice of the community. Further work examining the composition of the GCMU has been published elsewhere.

It is also worth noting that even though respondents were blind to the official outcomes, such as district-wide skilled birth attendance rates, many respondents perceived the intervention to have had positive outcomes on maternal mortality and have been beneficial to the community. We hypothesize that the sense of ownership may be stronger if the intervention is perceived to have a positive impact on maternal child health outcomes because community members would want to sustain these effects and the MWH model.

We had posited that the intervention would continue after the end of the external support period, in part due to the project’s goal of fostering community ownership. Respondent comments indicated that community ownership is connected to sustainability. Community members felt confident that the MWH was built for them and that they were able to use it. The community has a certain level of responsibility for the success of the MWH, but not necessarily sole responsibility. For the MWH to be sustainable, key roles need to be filled by different stakeholders. Specifically, the collaboration between the health facility and the GCMU is essential for the MWH to function well. They represent a linkage between those running the MWH and the community that must exist for a sense of community ownership to emerge and is viewed as an essential component of sustainability by respondents involved in the management of the MWH.

**Limitations**

There are several limitations with this analysis. First, our purposive sampling method, while allowing for a wider variety of opinions to be collected, may have resulted in over-representing the views of some groups such as women and farmers. However, our analysis of these perceptions was conducted across multiple time points and several stakeholder types, which strengthens our findings. Second, we had limited ability to explore some of our findings in greater depth. For example, our data do not allow us to know reasons for the lack of detailed mechanisms for conflict resolution or problem-solving within the MWH. Thirdly, project staff could have influenced the roles and responsibilities of stakeholders over time, through contact for project implementation and interviews themselves. These processes could have influenced the final outcome of how community ownership was expressed by respondents. However, because we did not use a pre-established definition for community ownership, we believe that our findings should not be overly impacted by this effect. Finally, our focus on community ownership emerged from our formative work, during which community members underscored the importance of this concept to the sustainability of a MWH intervention. We acknowledge that other populations may consider government ownership or other strategies as important routes toward sustainability.

**CONCLUSION**

Considering the need to ensure the sustainability of maternal health interventions, we found it essential to assess how stakeholders understood the concept of community ownership of an MWH model. Community ownership has long been an ill-defined concept with little evidence for how it is operationalized on the ground. We found variation in the definition of ownership by stakeholder type. While users described ownership as an ability to stay at the MWH and benefit from it, stakeholders involved in the operations of the MWH focused on the importance of the collaboration...
between the governance committee and the health facility staff, which respectively represent the community at-large and the larger health system. These definitions and perceptions are particularly important to consider when designing health interventions to ensure that their positive impact continues once external support is withdrawn.

Acknowledgments: The authors would like to thank the Zambian Ministry of Health at the national, provincial, and district levels as well as the traditional leadership of the relevant areas that approved and supported this study. We are deeply thankful for the community members and health system stakeholders who participated in the focus group discussions and in-depth interviews and shared their experiences, time, and perspectives. We would also like to thank the data collectors and study staff for their diligent collection of data and constant efforts. We would like to thank Viviane Sakanga, Kaluba Mataka, Denson Chongwe, Deophine Bwalya, Melvin Mwansa, and Edwin Tembo who were all critical in the implementation of the MWHs; the formation, training, and mentorship of the GCMU; and community sensitization efforts. We also appreciate Kathleen Lucile McGlasson who assisted with generating demographic information for Tables 1 and 2.

Competing interests: None declared.

REFERENCES


Bringing Greater Precision to Interactions Between Community Health Workers and Households to Improve Maternal and Newborn Health Outcomes in India

Peter Smittenaar, a B.M. Ramesh, b Mokshada Jain, a James Blanchard, b Hannah Kemp, a Elisabeth Engl, a Shajy Isaac, b,c John Anthony, b,c Ravi Prakash, b,c Vikas Gothwal, b,c Vasanthakumar Namasivayam, b Pankaj Kumar, d Sema K. Sgaier a,e,f

Key Findings
- Community health worker (CHW) presence, number and timing of visits, behavior change messaging strategies, and focus on specific household members for different behaviors associates with maternal and newborn care practices.
- Local sociocultural factors such as the decision dynamics of households and common false beliefs about neonatal care should inform how the CHW communicates.

Key Implications
- Program managers can use these insights to adapt the CHW training, incentives, and tools to achieve greater impact.
- Other CHW programs can use this approach to identify opportunities to improve the effectiveness of their workers in terms of communication, knowledge, and number and timing of visits.

ABSTRACT
Introduction: Community health workers (CHWs) play a key role in the health of mothers and newborns in low- and middle-income countries. However, it remains unclear by what actions and messages CHWs enable good outcomes and respectful care.
Methods: We collected a uniquely linked set of questions on behaviors, beliefs, and care pathways from recently delivered women (n=5,469), their husbands (n=3,064), mothers-in-law (n=3,626), and CHWs (accredited social health activists; n=1,052) in Uttar Pradesh, India. We used logistic regression to study associations between CHW actions and household behaviors during antenatal, delivery, and postnatal periods.
Results: Pregnant women who were visited earlier in pregnancy and who received multiple visits were more likely to perform recommended health behaviors including attending multiple checkups, consuming iron and folic acid tablets, and delivering in a health facility (ID), compared to women visited later or receiving fewer visits, respectively. Counseling the woman was associated with higher likelihood of attending 3+ checkups and consuming 100+ iron and folic acid tablets, whereas counseling the husband and mother-in-law was associated with higher rates of ID. Certain behavior change messages, such as the danger of complications, were associated with more checkups and ID, but were only used by 50%–80% of CHWs. During delivery, 57% of women had the CHW present, and their presence was associated with respectful care, early initiation of breastfeeding, and exclusive breastfeeding, but not with delayed bathing or clean cord care. The newborn was less likely to receive delayed bathing if the CHW incorrectly believed that newborns could be bathed soon after birth (which is believed by 30% of CHWs). CHW presence was associated with health behaviors more strongly for home than facility deliveries. Home visits after delivery were associated with higher rates of clean cord care and exclusive breastfeeding, and the newborn was less likely to receive delayed bathing if the CHW incorrectly believed that newborns could be bathed soon after birth (which is believed by 30% of CHWs). CHW presence was associated with health behaviors more strongly for home than facility deliveries. Home visits after delivery were associated with higher rates of clean cord care and exclusive breastfeeding, and the newborn was less likely to receive delayed bathing if the CHW incorrectly believed that newborns could be bathed soon after birth (which is believed by 30% of CHWs). CHW presence was associated with health behaviors more strongly for home than facility deliveries.
Conclusion: We identified potential ways in which CHW impact could be improved, specifically by emphasizing the importance of home visits, which household members are targeted during these visits, and what messages are shared. Achieving this change will require training CHWs in counseling and behavior change and providing supervision and modern tools such as apps that can help the CHW keep track of her beneficiaries, suggest behavior change strategies, and direct attention to households that stand to gain the most from support.
**INTRODUCTION**

In low- and middle-income countries, community health workers (CHWs) provide basic but lifesaving support for those who have little access to formal health care, especially for reproductive, maternal, neonatal, and child health services. Their impact on health outcomes has been documented and demonstrates considerable further potential.

One barrier to realizing this potential is a lack of understanding about why CHWs are effective in achieving some health outcomes, but less so in others. Most studies of CHWs are descriptive, for example, capturing CHWs’ level of clinical knowledge or the number of visits they make. Explanatory studies of the mechanisms by which CHWs achieve impact (or fail to) are needed. Such insights can be used to guide the development of CHW programs.

We report a detailed quantitative study of the interactions between household members and their CHW during pregnancy, delivery, and early postnatal care. The CHWs are part of India’s accredited social health activist (ASHA) program—the world’s largest CHW program, which has recruited about 1 million volunteers. ASHAs are trained by the government’s National Rural Health Mission and receive financial incentives for their work. Each ASHA is responsible for guiding about 1,000 people in her community through interactions with health services, providing basic medical and contraceptive supplies, and educating them on basic health topics.

In practice, ASHAs divide their time between accompanying people (primarily women) to health facilities, documenting their work, supporting regular village health and nutrition events, and visiting families in their homes.

A recent review of research on the ASHA program suggested an overall positive impact, but most studies showed mixed results in terms of performance and outcomes. This is in part explained by the constraints of the health system in which the ASHA operates, but also by the complexities of navigating decision dynamics in rural Indian households. Upon marrying, most women move in with their husband’s family, meaning that other family members often play decisive roles in certain health-related decisions. In Uttar Pradesh, where this study was performed, women rarely make maternal and child health decisions on their own. Rather, the husband and mother-in-law (MIL) play a major role. The husband’s knowledge about ANC practices and delivering in a health facility and the MIL’s knowledge about early initiation of breastfeeding (EIBF) practices are associated with whether the household performs these practices. Our qualitative work suggests men are most likely to be involved when decisions have financial repercussions, such as whether to deliver in a facility, whereas the MIL is deferred to for decisions around home care, such as breastfeeding or cord care (unpublished results). Social norms further impinge on these decisions, with norms themselves constantly evolving. Altogether, this leaves the ASHA with difficult decisions on who to target in the household to effect behavior change toward recommended health behaviors, and the optimal target is not necessarily the primary decision maker if this individual is not receptive to ASHA advice.

We designed a novel approach to study the mechanisms of CHW impact around pregnancy, childbirth, and early postnatal care. We collected detailed questionnaire data from recently delivered women, their husbands, MILs, and ASHAs. Understanding perspectives of each of the decision makers in the household allowed us to identify ways to make ASHA home visits more impactful. We set out to answer the following questions: (1) What are the rates of recommended behaviors? (2) What support are women receiving from their ASHA? (3) Is there an association between the number and timing of ASHA home visits or ASHA’s presence at birth and uptake of recommended health behaviors? (4) What types of behavior change messages are favored by ASHAs during the antenatal period, and is household behavior associated with favored messages? (5) To what extent is uptake of recommended behavior associated with whether the woman, husband, and mother-in-law are counseled by the ASHA? For this last question, we did not have strong hypotheses as to whom the ASHA should target, given that impact is a function of the ASHA’s ability to talk to the respective household member, receptiveness to advice by the recipient, and the recipient’s power in the household. We also examine the extent to which ASHA presence at birth is associated with respectful care, which is often compromised in low-resource settings.

Our goal is to stimulate reflection on the impact of a particular cadre of CHWs, especially how such impact is achieved, to generate lessons that may be relevant for other CHW programs with regard to training, supervision, and tools.

**METHODS**

**Sampling Household Members and CHWs**

The primary sampling unit was the ASHA catchment area, which usually comprises 1 village.
all rural catchment areas in India’s Uttar Pradesh state (UP), 1,575 were selected pseudorandomly, with higher likelihood of sampling assigned to those previously identified by the UP government as high-priority for interventions. The final sample of 1,575 areas had an average population of 1,064 (standard deviation [SD]=344) people as reported by the ASHA. All households with a living woman who had given birth in the past 2 months (including perinatal death cases) were approached for the survey. Between September 2017 and January 2018, 6,078 interviews of women were conducted across 1,514 ASHA catchment areas (no recently delivered women were identified in 61 catchment areas). Of these, 609 interviews were excluded because they were conducted in the first week after birth (thus did not yet have data on their postnatal behaviors), the women declined to consent, they were duplicates of later interviews (e.g., if the interview was initiated but then rescheduled), or they were discontinued midway through the survey. The final sample analyzed was 5,469 women across 1,499 catchment areas (76 of 1,575 catchment areas had no women included and were therefore excluded).

One ASHA was assigned to each catchment area. For most (n=5,278) of the final sample of 5,469 women, their ASHA was also surveyed if the catchment area had an ASHA assigned and she could be contacted (of 1,575 catchment areas, 1,502 ASHAs completed an interview, of whom 1,435 had at least 1 woman from their area interviewed). This approach allowed us to understand how traits, beliefs, experiences, behaviors, and other characteristics of the ASHA and woman were associated at the individual level.

Data were collected through face-to-face interviews at the homes of women and ASHAs. Details of the questionnaires, including design and content, are described in Supplement 1. Each respondent provided informed oral consent at the time of screening and written consent before the full interview. Interviewers were trained on sensitive topics and ethics of conducting interviews. The study was approved by Sigma Institutional Review Board (approval number 10032/IRB/D/16-17, New Delhi, India).

Public Involvement

There was no public involvement element. Before launching the main study, we conducted a pilot test where we interviewed a handful of recently delivered women, their husbands, MILs and ASHAs to learn from their experiences before, during, and after childbirth. Written consent was obtained and the information provided during the interview was treated as confidential. Based on the pilot, we refined the survey to make it simpler for both interviewers and respondents. The public was not further involved in the study nor will they be involved in dissemination of the results, which will be shared with health programs and the government.

Variable Selection and Statistical Analysis

We aimed to cover key actions taken by the ASHA, as well as key health behaviors and outcomes that are important for households. From the household survey, we selected 9 outcomes known to be important for maternal and neonatal health15,16: (1) attending 3 or more antenatal checkups at a monthly village health and nutrition day or at a medical facility; (2) taking at least 100 iron and folic acid (IFA) tablets over the course of pregnancy; (3) having an institutional delivery (ID) at a public or private facility; (4) receiving respectful care with no experience of physical or verbal mistreatment from staff; (5) staying in the facility for at least 24 hours after delivery for women without cesarean delivery only; (6) initiating breastfeeding within 1 hour of birth (EIBF) (for women without cesarean delivery only); (7) promoting exclusive breastfeeding (EBF); (8) applying nothing to the umbilical cord stump that might cause infection (clean cord care); and (9) delaying bathing of the baby by at least 72 hours. Although the 72-hour delay is longer than usually selected in studies,15 it is the period advised by the UP government.

We also defined 8 ASHA actions, as reported by the recently delivered woman, that had the potential to impact women’s health behaviors16: (1) whether any antenatal home visit was made; (2) number of antenatal home visits (for the subgroup with 1+ visits only); (3) month of pregnancy of the first home visit (for the subgroup with 1+ visits only); (4) whether the ASHA was present for any part of delivery, irrespective of the place of delivery; (5) how many hours the ASHA was present at delivery (for the subgroup that had an ASHA present only); (6) how many hours the ASHA was present immediately after birth minus the hours present before birth (for the subgroup with an ASHA present only); (7) whether any postnatal visit was made in the first week after birth; and (8) number of postnatal home visits in the first week after birth (for the subgroup with 1+ postnatal visits only).
Statistical Analysis: Logistic Regression
We used logistic regression to test for associations between ASHA actions and household behaviors, both as reported by the household. Each action-outcome association had its own specific set of control variables as well as a particular (sub)group of participants included, so a separate regression was performed for each association (Supplement 2 includes discussion of multiple comparisons). Nonetheless, each regression presented here included a set of demographic control variables captured from the woman (Supplement 3, Table): religion, caste, parity, years of education, age of woman, electricity available in house as proxy for wealth, and household type (nuclear versus multigenerational). The control variables were selected based on authors’ knowledge of the context, existing literature, and availability in the survey.

We implemented unweighted regressions in the R software package.17 To visualize the relevant relationships in graphs, we transformed the predictors to a categorical variable and performed the regressions again so we could calculate estimated marginal proportions corrected for all covariates.18 No adjustments were made for clustering of the data, as every ASHA on average only covered 4 recently delivered women in the survey sample.

We identified each ASHA’s “preferred message” for convincing women to attend checkups and go for ID, respectively, from the ASHA survey (Supplement 1). Household members and ASHAs were linked by a unique ASHA catchment identifier that was recorded in both surveys. We then regressed each woman’s attendance of 3+ checkups and ID, respectively, onto the preferred message of the ASHA in their catchment area. To partially control for features of the ASHA that might correlate with their preferred message, we not only used the set of covariates used for the main analyses but also added covariates derived from the ASHA questionnaire: demographics, working hours, tenure, earnings, and number of familiar messaging strategies.

RESULTS
The demographics of the women, husbands, and MILs are described in Supplement 3.

Rates of Recommended Health Behaviors Fell Below Target
We compared observed rates to target rates set by the government for the end of 2019, where available (Technical Support Unit, Uttar Pradesh, Phase 2 targets). Percentages were weighted for oversampling of government interest areas to ensure they are representative of the state of Uttar Pradesh, but n were not (denominator varied in part due to respondents could answer they did not remember what precisely happened). Only 49% (2,625/5,467) of pregnant women reported attending 3+ checkups (no target available), 14% (782/5,469) reported having taken the recommended 100 IFA tablets (target: 50%), 82% (4,428/5,469) of women had an ID (target: 80%), and 21% (817/3,898) of women reported having remained in hospital for at least 24 hours of birth (target: 80% stay for 48+ hours). In the postnatal care phase, 58% (2,823/4,925) of women reported EIBF (target: 80%); 63% (3,453/5,469) EBF (target: 80% for full 6 months); 51% (2,554/5,222) reported having first bathed their newborns 72+ hours after delivery (target: 90%); and 18% (915/5,268) reported not applying anything to the umbilical cord stump of their newborn (target: 90%).

ASHA Support Varied Across the Maternal and Newborn Care Pathway
The ASHA is expected to visit each woman at least 3 times during pregnancy. Most households (84%) reported at least 1 such visit (Figure 1A). For households receiving 1 visit, the ASHA visited relatively often (median 4 times across pregnancy) and early in pregnancy (at month 3.5 on average (Figure 1B); guidelines recommend the ASHA visit as soon as the woman is aware of the pregnancy. When the woman went into labor, the ASHA was incentivized to accompany her to a public facility or be present for home birth. The ASHA was present for 83% of deliveries that happened in a public facility (Figure 1C). However, the ASHA was largely absent from deliveries at home (present for 14% of deliveries) and in private facilities (present for 23% of deliveries); overall 57% of women reported that the ASHA was present at the delivery (Figure 1C). When the ASHA was present for delivery, she would often be there for several hours before and several hours after delivery (Figure 1D). However, sometimes the ASHAs were present for 16+ hours before delivery (7% of cases where ASHA was present) or 16+ hours after delivery (8% of cases where ASHA was present [Figure 1D]). Only 31% of women reported having received 2 visits or more in the first week after birth, and 30% reported having received no home visit at all over the first week after birth.

ASHA Work Hours Were Skewed Toward Hospital Visits and Paperwork
Time expenditure was determined by asking the ASHA how frequently she performed a particular
task (e.g., “accompanied woman to hospital” or “performed census”), and how long it took on average. Adding up her reproductive, maternal, and neonatal health activities and related administrative tasks (but excluding additional duties on other health topics), the average ASHA reported 65.7 hours per month on this work (Supplement 4, Supplemental Figure 1 has a visual representation). Her monthly work consisted of:

- Accompanying women to hospital for delivery or checkups (22 hours; average ASHA spent 33.5% of her time on this task)
- Documentation (19 hours; 28.9%)
- Attending Village Health and Nutrition Day (6.2 hours; 9.4%)
- Home visits to women after birth (5.8 hours; 8.8%)
- Home visits to pregnant women (4.2 hours; 6.4%)
- Home visits to couples eligible for family planning (3.4 hours; 5.2%)
- Village population census (2.6 hours; 4.0%)
- Home visits for child immunization (2.5 hours; 3.8%)

This showed that the bulk of the ASHA’s time was spent on accompanying women on visits to health facilities and on paperwork, with home visits comprising only 24.2% of her time spent on maternal, neonatal, and child health.
Health Behaviors Were Associated With Timing and Number of ASHA Visits and Preferred Messaging

**ASHA Support to Pregnant Women**

**Quantity and Timing of Antenatal Home Visits Was Associated With Health Behaviors.** Antenatal home visits had substantial positive associations with health behaviors (Table and Figure 2A). Receiving at least 1 visit was strongly associated with attending 3+ checkups (adjusted odds ratio [aOR]=2.22; 95% confidence interval [CI]=1.90, 2.59), ID (aOR=2.20; 95% CI=1.85, 2.61), and delayed bathing (aOR=1.55; 95% CI=1.30, 1.86). More visits further increased these odds (a positive “dose-response” effect, though we cannot say whether the home visit preceded check-ups or IFA consumption), with the benefit of additional visits tapering off around 4–6 visits: (Figure 2A); 3+ checkups aOR=1.14; 95% CI=1.11, 1.17); ID (aOR=1.09; 95% CI=1.05, 1.14); and delayed bathing (aOR=1.08; 95% CI=1.04, 1.11). Number of home visits was also associated with IFA consumption (aOR=1.05; 95% CI=1.01, 1.09) and cord care (aOR=1.07; 95% CI=1.03, 1.12), though these associations were relatively weak (Figure 2A). In contrast, number of antenatal home visits was not associated with duration of facility stay after birth (aOR for receiving at least 1 visit=0.98; 95% CI=0.78, 1.23), EBF (aOR=0.99; 95% CI=0.84, 1.18), and only weakly if at all with EIBF (aOR=1.21; 95% CI=1.00, 1.45).

We observed that earlier first visits by the ASHA increased the likelihood that the woman attended 3+ checkups (aOR=1.10; 95% CI=1.06, 1.15) and consumed 100+ IFA tablets (aOR=1.07; 95% CI=1.01, 1.13) (Table and Figure 2B; included a covariate for total number of antenatal home visits). However, no significant association was found with ID (aOR=1.05; 95% CI=0.99, 1.10) (Table).

**Targeting Advice to the Right Decision Maker Mattered During Antenatal Visits.** To better understand the mechanisms underlying the associations from Figure 2, we analyzed the targeting of household members with messages. The recently delivered woman, her husband, and her MIL each reported whether the ASHA advised them on checkups, IFA, and ID, respectively. We tested whether such advice was associated with behavior, noting that for behavior to change the ASHA needed to know how to communicate well to each particular household member and the recipient needed to have the required influence in the household. Attendance of 3+ checkups was more likely in households where the ASHA had advised the woman on antenatal checkups (Figure 3A): aOR for counseling of mother (independent variable) on 3+ checkups (dependent variable)=1.47; 95% CI=1.12, 1.93). Similarly, consumption of 100+ IFA tablets was more likely in households where the ASHA had counseled the mother on IFA tablets; women were equally likely to receive these tablets from either an ASHA or auxiliary nurse midwife (Figure 3A; aOR=1.48; 95% CI=1.08, 2.05). Counseling the husband or MIL on checkups or IFA, respectively, had no association with likelihood of performing such behaviors (Figure 3A). In contrast, advice to go for ID was associated with increased rates of ID only if the husband (aOR=2.16; 95% CI=1.66, 2.83; P<.001) and MIL (aOR=1.52; 95% CI=1.18, 1.95; P=.001) received it (Figure 3A). For each of the 3 behaviors, the pregnant woman was most likely to report having been counseled, followed by the MIL, and finally the husband (Supplemental Figure 2, Supplement 5).

ASHAs Lacked Awareness of Effective Behavior Change Messages. To investigate the type of message most likely to be associated with a target behavior, we asked all ASHAs what messages they used when trying to convince families to go for checkups and ID (Figure 3B–C). Financial arguments were most common (e.g., “Checkups are free”), whereas social norm arguments (e.g., “It’s common practice”) were least often reported. For each type of message, 50%–80% of ASHAs reported using them. The most commonly reported message types were not necessarily those most strongly associated with the target behaviors. The 2 least-reported message types for checkups—saying that checkups were now common and that doctors recommended them—were associated with the highest rate of checkup attendance (Figure 3D; P values against reference of “checkups are free”: checkups are common, P=.05; doctors advocate checkups, P=.08; detected complications early, P<.001; stories of own experience, P=.04; all other P>.10). Conversely, the commonly reported message that “checkups are free” was associated with the lowest rate of 3+ checkups. For ID, messages about complications were most strongly associated with ID, whereas the least reported “ID is common” was associated with the lowest rates of ID (Figure 3E; P values against reference of “ID incentive”: complications are dangerous, P=.02; stories of dangerous births, P=.02; ID is common, P=.07; all other P≥.1).
ASHA Presence at Birth

ASHA Presence Was Associated With EIBF and EBF. ASHAs encouraged families to go to a hospital for birth and often would stay for hours (Figure 1D). ASHA presence at delivery was associated with a 7-percentage-point higher likelihood of EIBF (aOR=1.32; 95% CI=1.12, 1.56) and a 6-percentage-point higher likelihood of EBF (aOR=1.24; 95% CI=1.06, 1.45) (Table and Figure 4). Although women generally reported high levels of respectful care by staff, ASHA presence during delivery was associated with a 3.6-percentage-point higher likelihood of women receiving respectful care (aOR=1.55; 95% CI=1.14, 2.07) (Figures 4A

<table>
<thead>
<tr>
<th>CHW action</th>
<th>Any Antenatal Home Visit</th>
<th>Number of Antenatal Home Visits</th>
<th>Timing of First Antenatal Home Visit</th>
<th>Present for Any Part of Delivery</th>
<th>Hours Present Before Delivery</th>
<th>Pre- vs Postpartum Hours (After Before)</th>
<th>Any Home Visit in First Week After Delivery</th>
<th>Number of Home Visits in First Week After Delivery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antenatal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Received 3+ checkups</td>
<td>2.22</td>
<td>1.14</td>
<td>1.10</td>
<td>1.32</td>
<td>1.01</td>
<td>1.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.90, 2.59)</td>
<td>(1.11, 1.17)</td>
<td>(1.06, 1.15)</td>
<td>(95% CI=1.12, 1.56)</td>
<td>(0.99, 1.03)</td>
<td>(0.99, 1.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consume 100+ IFA</td>
<td>1.11</td>
<td>1.05</td>
<td>1.07</td>
<td>1.24</td>
<td>1.01</td>
<td>1.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.9, 1.37)</td>
<td>(1.01, 1.09)</td>
<td>(1.01, 1.13)</td>
<td>(95% CI=1.06, 1.45)</td>
<td>(0.99, 1.04)</td>
<td>(0.99, 1.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delivery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delivered in facility</td>
<td>2.20</td>
<td>1.09</td>
<td>1.05</td>
<td>1.21</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.85, 2.61)</td>
<td>(1.05, 1.14)</td>
<td>(0.99, 1.10)</td>
<td>(95% CI=1.00, 1.45)</td>
<td>(0.98, 1.01)</td>
<td>(0.98, 1.01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respectful care</td>
<td></td>
<td></td>
<td></td>
<td>1.55</td>
<td>1.01</td>
<td>1.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1.14, 2.07)</td>
<td>(0.99, 1.03)</td>
<td>(0.99, 1.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stayed at facility for 24+ hours</td>
<td>0.98</td>
<td>1.01</td>
<td>1.02</td>
<td>1.55</td>
<td>1.01</td>
<td>1.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.78, 1.23)</td>
<td>(0.97, 1.05)</td>
<td>(0.82, 1.27)</td>
<td>(95% CI=1.14, 2.07)</td>
<td>(0.99, 1.04)</td>
<td>(0.99, 1.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postnatal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early initiation breastfeeding</td>
<td>1.21</td>
<td>0.99</td>
<td>1.32</td>
<td>1.21</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.00, 1.45)</td>
<td>(0.96, 1.02)</td>
<td>(1.12, 1.56)</td>
<td>(95% CI=1.00, 1.45)</td>
<td>(0.99, 1.01)</td>
<td>(0.98, 1.01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delayed bathing 72+ hours</td>
<td>1.55</td>
<td>1.08</td>
<td>0.95</td>
<td>1.21</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.30, 1.86)</td>
<td>(1.04, 1.11)</td>
<td>(0.81, 1.12)</td>
<td>(95% CI=1.08, 1.45)</td>
<td>(1.00, 1.01)</td>
<td>(0.99, 1.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exclusively breastfed</td>
<td>0.99</td>
<td>1.01</td>
<td>1.24</td>
<td>1.44</td>
<td>1.00</td>
<td>1.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.84, 1.18)</td>
<td>(0.98, 1.05)</td>
<td>(1.09, 1.45)</td>
<td>(95% CI=1.25, 1.66)</td>
<td>(0.99, 1.02)</td>
<td>(0.99, 1.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clean cord care</td>
<td>1.00</td>
<td>1.07</td>
<td>1.19</td>
<td>1.56</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.78, 1.29)</td>
<td>(1.03, 1.12)</td>
<td>(0.96, 1.49)</td>
<td>(95% CI=1.14, 1.73)</td>
<td>(0.98, 1.01)</td>
<td>(0.98, 1.01)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CHW, community health worker; CI, confidence interval; IFA, iron and folic acid.

A women with 1+ antenatal home visit only.

B Women with community health worker presence only.

C Women with 1+ postnatal home visit within first week after birth only.

D An adjusted odds ratio (95% CI) greater than 1 means the household behavior was more likely. Each aOR is from a separate regression with appropriate covariates depending on the relationship being estimated, including corrections for demographics, parity, location of delivery, and other CHW actions.
and 4D). ASHA presence was not found to be associated with clean cord care (aOR=1.19; 95% CI=0.96, 1.49) or delayed bathing (aOR=0.95; 95% CI=0.81, 1.12). We also did not observe an association between the duration of ASHA presence and any of the 4 postnatal care behavior outcomes or between her presence before versus after birth and any of the outcomes (see Table for aOR values).

**ASHA Presence Had a Protective Effect for Home Deliveries.** ASHA presence at home deliveries was associated with higher likelihood of EIBF (+12 percentage points, \(P=.01\)), EBF (+10 percentage points, \(P=.04\)), and clean cord care (+5 percentage points, \(P=.05\)). In contrast, at public facilities, the only significant association was with EBF (+5 percentage points, \(P=.04\)). At private facilities, the only significant association was with EIBF (+10 percentage points, \(P=.04\)). The association between ASHA presence and cord care was significantly greater for home deliveries than public (\(P=.03\)) or private (\(P=.04\)) deliveries. Similarly, respectful care was more likely when the CHW was present for public-facility deliveries compared to private-facility deliveries (\(P=.001\)). In contrast, none of the other behaviors (EBF, EIBF, or delayed bathing) showed a stronger association with ASHA presence at home compared with birth in facilities. These marginally significant findings suggest that ASHA presence at home births might be particularly impactful.

**Incorrect ASHA Knowledge Was Associated with Early Bathing.** Thirty percent of ASHAs incorrectly believe bathing immediately after birth did not put the newborn at risk. We found that if the ASHA had incorrect knowledge on delayed bathing, her presence at birth was associated with lower likelihood of delayed bathing (Figure 4C; presence-knowledge interaction aOR=2.07; 95% CI=1.33, 3.22; \(P<.001\)). The marginal proportions suggested that incorrect knowledge of ASHAs regarding recommended delayed bathing practices might be encouraging incorrect behaviors (Figure 4C).

**ASHA Actions in Postnatal Period**

**ASHA Visits Had a Protective Effect on Clean Cord Care and Exclusive Breastfeeding.** Receiving 1 or more ASHA visits (compared to receiving none) in the first week after birth was associated with both higher likelihood of EBF (aOR=1.44; 95% CI=1.25, 1.66) and clean cord care (aOR=1.40; 95% CI=1.14, 1.73). For women reporting they had received at least 1 visit, additional visits further improved the odds of those behaviors: EBF (aOR=1.21; 95% CI=1.08, 1.36) and clean cord care (aOR=1.56; 95% CI=1.37, 1.77) (Table and Figure 5A). Those reporting 3+ visits

---

**FIGURE 2.** Estimated Relationships Between Health Behaviors and (A) Number of CHW Home Visits and (B) Timing of First Home Visit

Abbreviations: ANC, antenatal care; CHW, community health worker; IFA, iron and folic acid. Only significant associations are shown, see Table 1 for all associations. All means are estimated marginal proportions adjusted for covariates.

\(^{a} n=5,438 \text{ women.}^{b} n=4,541 \text{ women, excluding those who received no visit.}\)
had substantially better odds of performing recommended behaviors (Figure 5A).

**Targeting MIL Positively Was Associated With Exclusive Breastfeeding.** Advice from the ASHA to the MIL was associated with substantially higher rates of EBF (aOR=1.54; 95% CI=1.24, 1.92; P<.001) (Figure 5B, top), whereas the likelihood of clean cord care was unaffected by the ASHA’s advice (Figure 5B, bottom). The data were insufficiently granular to tease apart the relationship between timing of visits and when the decisions were actually made in the household. As with antenatal counseling, the recently delivered woman was most likely to report having been counseled, followed by the MIL and finally the husband (Supplemental Figure 2, Supplement 5).

**DISCUSSION**

This study analyzed a large cross-sectional data set from women who had delivered a newborn in the preceding 2 months in UP, India, their husbands, MILs, and ASHAs. Having such rich data on each household allowed us to explore possible relationships between ASHA activities and target behaviors and gain insights to improve ASHA performance. As with earlier studies,19–22 we found that ASHA home visits were associated with a range of target behaviors. The study also offers insights into the influence of ASHA knowledge, timing of visits, messaging strategies, and decision-maker targeting on care practices. Such insights can enable program managers to further adapt the ASHA program to achieve greater impact. Other CHW programs can...
also learn from our findings when considering what guidance to give their workers on communication, knowledge, and number and timing of visits.

In the survey, two-thirds of pregnant women reported having received 3 antenatal home visits from their ASHA, though we observed better health behaviors up to 6 or 7 antenatal home visits. There are no guidelines for ASHAs on how many home visits to make and when,\(^2\) though the 2016 World Health Organization recommendation is that a woman attends 8 antenatal clinic visits during pregnancy.\(^1\) In resource-limited settings such as those in UP, CHW home visits might have to supplement a smaller number of clinic visits. In the first week after birth, only 31\% of women receive 2+ visits, and we observed better health behaviors for up to 3+ postnatal visits in the first week after birth. ASHAs spend considerably less time visiting households than the 34–65 hours monthly suggested by guidelines.\(^2\) Therefore, it seems that increasing the number of home visits would be beneficial. Improving the content of each visit (e.g., through stronger communication strategies) might help ASHAs achieve behavior change more readily, thus reducing the need for additional home visits. Either way, we recommend clearly communicating a guideline for the number of antenatal home visits ASHAs should make, which we were unable to

**FIGURE 4.** (A) Relationship Between Outcomes and CHW Presence at Birth Irrespective of Location of Delivery\(^a\); (B) Relationship Between Outcomes and CHW Presence at Birth by Location of Delivery\(^b\); (C) Association of CHW’s Awareness of Delayed Bathing and Practice of Delayed Bathing When CHW Was Absent or Present\(^c\); (D) Rate of Different Forms of Poor Treatment at Public Facility Deliveries When the CHW Was Absent or Present\(^d\)

Abbreviations: BF, breastfeeding; CHW, community health worker.

\(^a\) N=5,240 women.  
\(^b\) N=984 women at home, 1068 at private, 3123 at public.  
\(^c\) N=4831 women.  
\(^d\) N=3244 women. Error bars represent 95\% CI.
identify in ASHA training materials. We also note the absence of a relationship between antenatal home visits and recommended breastfeeding behaviors. We have no data on counseling about postnatal care during antenatal home visits, but counseling on EIBF, EBF, and clean cord care—to both the woman and her MIL—should be a priority in the final few months of pregnancy. This is especially critical as several such postnatal behaviors are subject to strong social norms such as feeding newborns with jaggery (as opposed to exclusive breastfeeding) or applying substances to the cord that can cause infections. It is unclear to what extent the ASHA alone can overcome such norms, and most likely a government-driven effort is needed as was done to establish facility delivery as a norm.\textsuperscript{24}

Previous research often recommends that ASHAs improve their communication skills,\textsuperscript{10} but specific recommendations are lacking or based on small qualitative samples. ASHA training gives general communication tips but lacks specific recommendations to influence households on key behaviors and rarely mentions involving other stakeholders within the household. We found that many ASHAs did not use many behavior-change messages that could have been effective in driving higher rates of ANC visits and ID. They need to learn when to apply different messages and which individuals in the households should be counseled to achieve greatest impact (which may or may not be the primary decision maker). Our findings suggest that on average, the ASHA should target the pregnant woman for messaging about IFA and checkups, the husband and MIL for ID, and the MIL regarding postnatal care behaviors. Cultural barriers can inhibit the ASHA from talking to the husband or MIL,\textsuperscript{10} so supervision and peer-learning structures may support ASHAs in solving these challenges. A study on improving communication by the ASHA around pneumonia is under way,\textsuperscript{25} and our hope is that such initiatives will be expanded to maternal and neonatal domains. Our findings could also inform future editions of the ASHA training manuals, Sangini

---

**FIGURE 5.** (A) Association of Number of PNC Visits With Rate of Exclusive Breastfeeding and Clean Cord Care\textsuperscript{a}; (B) Increase in Health Behavior if Household Member Reported Having Talked to CHW About the Particular Behavior\textsuperscript{b}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure5}
\caption{(A) Association of Number of PNC Visits With Rate of Exclusive Breastfeeding and Clean Cord Care; (B) Increase in Health Behavior if Household Member Reported Having Talked to CHW About the Particular Behavior.}
\end{figure}

Abbreviations: BF, breastfeeding; CHW, community health worker; MIL, mother-in-law; PNC, postnatal care.

\textsuperscript{a} N=5,185 women.

\textsuperscript{b} N=2,057 households. Error bars represent 95\% confidence interval.
(supervisor) training, and technology-enabled job aids for ASHAs.

Potential guidelines that might result from synthesizing these antenatal results include: the ASHA should visit a woman at her home as soon as possible after learning she is pregnant and 4–6 times over the course of the pregnancy. During these visits, the ASHA should focus on the woman regarding IFA and antenatal checkups and use a range of messaging strategies such as emphasizing that checkups are now common, doctors recommend them, and that without checkups the woman is at risk of serious undetected complications. During these visits, she should also counsel the husband and MIL on the merits of ID, emphasizing the dangers of home delivery and the risk to life in case of complications. The ASHA should further advise the woman and MIL regarding neonatal care behaviors to increase the likelihood of recommended behaviors in the hours after birth. Although the details might be different for other CHW programs, similar analyses performed on other CHW cadres across the world could help identify opportunities for improvement.

The ASHA’s presence at births in public and private facilities was only weakly associated with targeted health behaviors and respectful care by staff (which many women in LMICs do not receive). There are several plausible mechanisms behind the observed associations: for example, adoption of recommended breastfeeding behavior may have been due to ASHA counseling, direct support, interaction with facility staff, or other reasons. In the case of respectful care, the ASHA did not wield direct power over the nurses or doctors at the facility. However, the ASHA was familiar with these individuals, was part of the health system, and could therefore have facilitated a positive relationship between community and facility staff. The duration of the ASHA’s presence at birth did not seem to affect any health behavior, throwing into question whether the many hours spent after accompanying the woman to hospital are an optimal use of her time. An important caveat is that we did not investigate other potentially positive effects of the ASHA’s presence at facilities, nor take into account the time spent at the facility waiting for a sign-off for her incentives or for transport home. Nevertheless, the reasons and potential benefits of her presence should be explored further to determine how useful it is for her to spend such long periods of time at facilities. The ASHA might use some of this time better making home visits to which she dedicates less than half as much time as facility visits and indeed less time than she spends on paperwork.

The ASHA’s influence seems more substantial at home deliveries, where the only other outside support usually comes from a traditional birth attendant. The disconnect between the ASHA’s training—which requires her presence for home births—and the low actual rate of presence can partly be explained by the fact that she receives no financial incentive for attending. A direct incentive would contradict the government’s drive for ID, but more creative ways of encouraging the ASHA’s presence could be considered (e.g., an incentive to accompany home-delivered babies for checkups at facilities soon after delivery, which would encourage an ongoing relationship between ASHA and household).

Past work has extensively studied CHW knowledge levels as a proxy for performance. We were able to show an association between ASHA lack of knowledge of delayed bathing and actual timing of first bathing by households. Therefore, we suggest that guidelines for ASHAs and their supervisors be more explicit about delayed bathing, since the current guidelines are ambiguous, advising only not to bathe the baby “immediately after birth.” The World Health Organization advocates a delay of at least 24 hours, and the government of UP sets 72 hours as a target.

These findings from UP add to a substantial body of evidence across the developing world suggesting that CHWs can play an important role in safeguarding the health of women and their newborns, especially those in the poorest parts of society. Their position on the frontlines provides a bridge especially for the rural poor into the formal healthcare system. However, it is often hard to understand why particular actions predict good health outcomes or why certain actions seem ineffective. Our findings show that local sociocultural factors such as the decision dynamics of households and common false beliefs about neonatal care should inform how the CHW communicates. It is noteworthy that a review found that of 31 training resource packages for CHWs across low- and middle-income countries, more than half contained only materials that were not tailored to the local context. Other CHW programs should consider the extent to which our findings might apply to their context and update their guidelines and incentive structures to reflect the optimal number and timing of visits, how CHWs are spending their time, and communication strategies. If resources allow, CHW programs should perform their own...
though not eliminating confounding. Empirical studies in the field could test for any causal nature to the associations uncovered here, for example, comparing 2 regimens of antenatal home visits with differing number of visits, or randomly assigning messages to ASHAs to use during home visits to test which ASHA actions are most likely to raise recommended health behavior rates. Second, the household behaviors are only partially under the control of the CHW. For example, a 24-hour facility stay after the birth, which was entirely unrelated to CHW actions, might be determined by the number of beds in a facility and availability of food. Similarly, the ability to get to facilities when labor starts depended on transport being available. ASHAs are only 1 piece of the puzzle of improving health and outcomes, and care should be taken not to shoulder ASHAs with responsibilities where they cannot exert control. In UP especially, ASHAs are often of a lower caste, meaning they might be restricted in what households they can visit and who they can counsel within the household. Others have not found a relationship between the mother’s caste and the number of ASHA visits received, nor was the relationship between ASHA support and health behaviors modified by caste. Nonetheless, given the known social dynamics around caste, further investigation potentially using the data collected here is warranted.

**CONCLUSION**

The mechanisms by which CHWs can contribute to better health for their communities are complex. Each household behavior faces its own set of contextual and perceptual drivers (a framework to comprehensively enumerate such factors can be found in Engl et al.), and progress hinges on understanding these. Strengthening CHW programs worldwide requires going beyond descriptive reports to understand the causal pathways by which the CHW achieves impact. As we continue to ask more of CHWs, we must provide them with the tools and skills to manage an ever-increasing diversity of tasks.

**Funding:** This study was funded by individual donor contributions to the Surgo Foundation.

**Competing interests:** None declared.

**REFERENCES**


Peer Reviewed

Received: January 9, 2020; Accepted: July 13, 2020; First Published Online: September 22, 2020


© Smittenaar et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are properly cited. To view a copy of the license, visit http://creativecommons.org/licenses/by/4.0/. When linking to this article, please use the following permanent link: https://doi.org/10.9745/GHSP-D-20-00027
Impact of Improved Biomass and Liquid Petroleum Gas Stoves on Birth Outcomes in Rural Nepal: Results of 2 Randomized Trials

Joanne Katz, a James M. Tielsch, b Subarna K. Khatry, c Laxman Shrestha, d Patrick Breysse, a Scott L. Zeger, a Naoko Kozuki, a William Checkley, e Steven C. LeClerq, a Luke C. Mullany a

Key Findings

- Two trials in rural southern Nepal reduced indoor air pollution through improved biomass or liquid petroleum gas stoves, but levels were still much higher than World Health Organization standards.
- Exposure of pregnant women to these lower air pollution levels was not associated with reduced incidence of low birth weight, preterm birth, and small for gestational age.

Key Implications

- Improved biomass stoves may not reduce indoor air pollution as much as is needed to have an impact on adverse birth outcomes.
- Emphasis should be placed on behavior to reduce use of biomass stoves when using improved ones for maximum reduction of pollution.

ABSTRACT

Background: Few randomized trials have assessed the impact of reducing household air pollution from biomass stoves on adverse birth outcomes in low-income countries.

Methods: Two sequential trials were conducted in rural low-lying Nepal. Trial 1 was a cluster-randomized step-wedge trial comparing traditional biomass stoves and improved biomass stoves vented with a chimney. Trial 2 was a parallel household-randomized trial comparing vented biomass stoves and liquid petroleum gas (LPG) stoves with a year’s supply of gas. Kitchen particulate matter of 2.5 μm or less (PM2.5) and carbon monoxide (CO) were assessed before and after stove installation. Prevalent and incident pregnancies were enrolled at baseline and throughout the trials. Birth anthropometry was compared across differing exposure times in pregnancy.

Results: In trial 1, the mean 20-hour kitchen PM2.5 concentration was reduced from 1380 μg/m³ to 936 μg/m³. Among infants born before the intervention, mean birth weight and gestational age were 2627 g (SD=443) and 38.8 weeks (SD=3.1), and 39% were low birth weight (LBW), 22% preterm, and 55% small for gestational age (SGA). Adverse birth outcomes were not significantly different with increasing exposure to improved stoves during pregnancy. In trial 2, the mean 20-hour PM2.5 concentration was 885 μg/m³ in households with vented biomass and 442 μg/m³ in those with LPG stoves. Mean birth weight was 2780 g (SD=427) and 2742 g (SD=431), among households with vented and LPG stoves, respectively. Respective percentages for LBW, SGA, and preterm were 23%, 13%, and 42% in the vented stove group and not statistically different from 31%, 17%, and 42% in the LPG group.

Conclusions: Improved biomass or LPG stoves did not reduce adverse birth outcomes. PM2.5 and CO following improved stove installation remained well above the World Health Organization indoor air standard of 25 μg/m³ or intermediate air quality guideline of 37.5 μg/m³. Trials that lower indoor air pollution further are needed.

INTRODUCTION

Low birth weight (LBW), comprising almost exclusively small for gestational age (SGA) and preterm births, is strongly associated with morbidity and mortality in infancy.1,2 Observational studies have shown associations between reported use of biomass fuel and these...
adverse birth outcomes but are subject to residual confounding even after adjustment for socioeconomic characteristics.\textsuperscript{3–9} One study showed an adjusted 43% increased risk of preterm birth and 21% increase in SGA with use of biomass fuel, but measures of exposure were limited to reported use of solid fuels.\textsuperscript{3} A meta-analysis of 19 studies found an 86.43 g reduction in birth weight and a 35% increase in LBW associated with biomass fuel exposure.\textsuperscript{9} Another meta-analysis of 5 studies showed that newborns in households using biomass fuels had a reduction in birth weight of 95.6 g and a 38% increase in LBW.\textsuperscript{7} Although multiple observational studies exist, only 2 randomized trials have examined the impact of reduced indoor air pollution on birth outcomes through the introduction of improved biomass or ethanol stoves.\textsuperscript{10,11} A trial in the highlands of Guatemala observed a 39% reduction in carbon monoxide (CO) levels, an increase in mean birth weight of 89 g, and a 26% lower rate of LBW among infants in households with improved biomass stoves compared with traditional ones.\textsuperscript{10} However, few pregnancies (n=266) were enrolled and impacts on birth outcomes were not statistically significant. A more recent trial from Ibadan, Nigeria, enrolled 324 pregnancies and found an 88 g higher mean birth weight and 40% reduced prevalence of preterm birth among infants born in households provided with ethanol stoves.\textsuperscript{11}

In 2010, it was estimated that 2.8 billion people used solid biomass fuels.\textsuperscript{12} Hence, if reductions in indoor air pollution exposures can be shown to have an impact on adverse birth outcomes, such interventions could potentially lower morbidity and mortality rates for a large number of newborns worldwide.

We report here the results of 2 randomized trials to reduce indoor air pollution caused by open burning of biomass fuel sources in the home and the effects on mean birth weight and gestational age, as well as LBW, SGA, and preterm births.

**METHODS**

**Population**

Two sequential trials were conducted in the low-lying plains of southern Nepal (Sarlahi District) contiguous with Bihar, northern India. This area is rural and has almost universal use of traditional mud brick biomass stoves (wood, dung, crop waste). The population consists primarily of subsistence farmers with low literacy rates. About half of women deliver at home and the prevalence of LBW, preterm, and SGA births is high.\textsuperscript{13,14} In trial 1, the study area consisted of eligible households in 4 village development committees (VDCs) in Sarlahi. Eligible households in 2 of these 4 VDCs were enrolled in trial 2. Trial 1 was conducted between March 2010 and August 2012, and trial 2 between March 2013 and March 2014.

**Study Design**

The study methods have been described in detail elsewhere.\textsuperscript{15} The aim of trial 1 was to examine whether lowering indoor air pollution by using an improved biomass stove could reduce acute lower respiratory illness in children and adverse birth outcomes among pregnant women. At the end of the first trial, indoor particulate matter (PM) was reduced by the improved stove but was still very high. Therefore, a second trial was designed to examine whether cleaner fuel, such as liquid petroleum gas (LPG), could further decrease the exposure and increase the likelihood of improved health outcomes.

For trial 2, because households from trial 1 and newly eligible households were enrolled and randomized, some pregnant women who were previously enrolled in trial 1 may have changed the type of stove they used, depending on the randomization in trial 2. A total of 126 women were still pregnant at the end of trial 1 and were eligible to enroll in trial 2.

Trial 1 was a randomized step-wedge community-based trial. Households were eligible if they did not have an LPG, electric, or improved vented biomass stove; had house materials that did not constitute a fire hazard with chimney installation; had at least 1 married woman 15–30 years of age or at least 1 child <36 months of age in the household at the start of the trial; and consented to participate. Women who were pregnant at the start of the trial were included and were assumed to have been exposed to traditional biomass stoves prior to enrollment in the trial. Over a 2-year period, these women were monitored for pregnancy with regular visits every 5 weeks. If women reported missing a period since the previous visit, project staff offered a pregnancy test; newly identified pregnant women were then followed until an outcome (miscarriage, live birth, or stillbirth). This process meant that for incident pregnancies, the recall time for date of last menstrual period was not very long, improving the quality of gestational age compared with asking about date of last menstrual period at delivery. The data collection for reproductive outcomes began with a 6-month trial of uniform...
surveillance in all enrolled households. This period was followed by a 12-month step-in period in which each household eventually had their open burning stove replaced with an improved version that had a chimney venting to the outside. The timing of the replacement from traditional biomass to improved biomass stove was randomized. Another 6-month period of surveillance for reproductive outcomes followed the 12 months of stove replacement. This design provided a similar time period and seasons before and after improved stove installation in the enrolled households (Figure 1).

In trial 2, households were eligible if they were in 2 of the 4 VDCs from trial 1, had at least 1 married woman 15–30 years of age or at least 1 child <24 months of age, and consented to participate. Study area households not enrolled in trial 1 were eligible for trial 2 if they met the same eligibility criteria as trial 1 (Figure 2). Households were individually randomized to use the improved biomass stove from trial 1 or to receive an LPG stove with a free year’s supply of gas. These households were followed for 1 year from the time of installation and birth outcomes within this time period were compared. Although this design was a traditional randomized trial, the length of time a pregnancy was exposed to the intervention varied.

**Randomization**

In trial 1, 51 clusters of 20–30 households comprising about 25–40 children <3 years of age were identified. This number of households was the approximate number in which a single installation team could install improved stoves in 1 month. The clusters were then grouped such that all enrolled households could have stoves installed over 12 months, and those groups were randomized to the timing of receipt of the intervention. The numbers 1 through 12 were written on slips of paper, and senior project staff selected a number from a hat. The numbers represented the month of the year in which the intervention would be provided for each of the 12 groups of clusters. In trial 2, households were randomized stratified on VDC and ward (a smaller administrative unit, with 9 in each VDC) using the sample procedure in STATA.

**Sample Size**

The sample size for both trials was based on the number of households needed to detect a difference in acute lower respiratory infection rates. The number of live births estimated to occur in trial 1 households was 2350 and 600 in trial 2. Given these live births, the detectable difference in birth weight was 57 g in trial 1 and 83 g in trial 2 with power of 80% and a type I error of 5%. Assuming 29% of infants would be born LBW (<2500 g) before the intervention (based on prior studies in this area), we had >80% power to detect a 20% relative reduction in LBW in trial 1, and 80% power to detect a 34% relative reduction in

FIGURE 1. Nepal Cookstove Intervention Trial, Phase 1 CONSORT Diagram
LBW in trial 2. Correlation of the outcomes within clusters was considered in the sample size calculations. Positive correlation increased the sample size required. The prevalence of adverse birth outcomes observed before the intervention in trial 1 was 39% LBW, 48% SGA, and 22% preterm. Trial 2 may have been underpowered because 9% fewer pregnancies were enrolled than the anticipated number of 600.

Interventions
In trial 1, the intervention was a 2-burner biomass stove with a chimney to vent smoke to the outside. This stove was manufactured by Envirofit International (Colorado Springs, CO, USA), model G-3300 with G-3355 2 pot attachment.15 The stoves were installed by a specially trained study team, and household members were trained in their use and maintenance. Stove monitors visited households weekly to encourage and record stove use.

In trial 2, households were randomly assigned to either continue using the same Envirofit stoves as in trial 1 or were provided an LPG 2-burner stove with gas tank. Study workers provided each household with 1 tank per month, which was estimated to provide gas for an average household of 5 for the 12-month duration of the trial. At the end of the study, households were also provided an additional month of fuel. Households assigned to continued use of the Envirofit stove were provided LPG stoves and a 1-month supply of gas at the end of the trial.

Exposure Measures
In trial 1, monitors to collect daily particulate matter of 2.5 μm or less (PM$_{2.5}$), CO, humidity and temperature were deployed in each household at least once prior to and once post installation of the vented stove. On average, indoor air pollution was measured 6 months prior to and 6 months after stove installation. Measurements were made every 10 seconds over an average of a 21.7-hour period (from approximately 3:00 pm through 12:00 pm the next day). Lunch is not a usual meal in Nepal, so this monitoring period likely covered nearly all cooking events in a day. The package of instruments was placed approximately 1.5 m above the floor and about 1 m in front of the stove to best represent exposure of the person cooking. Details of the indoor air pollution monitoring and impact have been described elsewhere.16

We measured CO and PM$_{2.5}$. We included CO measurements because we hypothesized that CO might be directly associated with adverse birth outcomes through the mechanism of oxygen displacement. In trial 2, the measurement taken post installation of the Envirofit stove in trial 1 was used as the PM$_{2.5}$ exposure measure for trial 2 households randomized to retain the Envirofit stove. Exposure measures were obtained in LPG households post installation (between 1 and 12 months later, average 6 months) in trial 2 as a comparison. Temperature and humidity were measured with a HOBO U10 Temperature and Humidity (TH) Data Logger (Onset Computer Corporation, Pocasset, MA). CO was measured with the LASCAR CO data logger (EL-USB-CO300, unpublished results).
Trials of Reduced Indoor Air Pollution to Improve Birth Outcomes

Bound of the weekly published data. Primary growth population distribution using the upper weights fell below the 10th percentile of the inter-completed gestation. SGA births were defined as those births occurring before 37 weeks completed from date of last menstrual period (collected every 5th week by local female study workers visiting the homes of enrolled women) to birth. Gestational ages that fell outside a feasible range of 24 to <43 weeks were treated as missing. Preterm births were those occurring before 37 weeks completed gestation. SGA births were defined as those whose sex- and gestational-age-specific birth weights fell below the 10th percentile of the inter-growth population distribution using the upper bound of the weekly published data. Primary outcomes compared between stove interventions were mean birth weight, mean gestational age, LBW, preterm birth, and SGA.

Statistical Analysis

For trial 1, descriptive demographic and socioeconomic characteristics were provided at baseline, prior to any stov installation. Given the step-wedge design, these characteristics were assumed not to change from pre to post installation. For trial 2, these same characteristics were compared between households that were randomized to the LPG intervention and those randomized to the improved biomass stoves to assess comparability between the 2 groups.

To characterize the impact of the vented biomass and LPG stoves, the mean daily PM$_{2.5}$ in micrograms per cubic meter and CO in parts per million (ppm) were compared with traditional stove measures in both trials. In addition, we examined the number of hours the kitchen measures of PM$_{2.5}$ were above 100 g/m$^3$ and above 1000 g/m$^3$ and CO measures were above 6 ppm and above 9 ppm and compared these measures of exposure between intervention groups. We chose these cut points for PM$_{2.5}$ and for CO because we wanted to describe average daily exposure and an ambient level in the absence of a cooking event. The median daily exposure to PM$_{2.5}$ was 1070 g/m$^3$ and the mean was 1380 g/m$^3$. Therefore, 1000 g/m$^3$ was chosen as a cut point. For CO, the median was 8.1 ppm and the mean was 11.0 ppm. We also used a statistical approach to determining when a cooking event was occurring (spikes in exposure could be seen). The lower cut point thus examined average ambient kitchen levels at times when cooking was not occurring. This cut point corresponded to PM$_{2.5}$ and CO at about 100 g/m$^3$ and 6 ppm, respectively.

In trial 1, women could have experienced various lengths of exposure to the intervention stove while pregnant, ranging from no exposure to exposure for the entire pregnancy, given data collection occurred over 2 years. This exposure calculation applied to both existing (prevalent) and new (incident) pregnancies. Exposure for incident pregnancies was based on when in pregnancy the improved stove was installed. Given that the length of time exposed to the improved stove could affect the impact of the intervention, we stratified the effect of the intervention by time of exposure in utero (no exposure in pregnancy, >0 to <1/3, 1/3 to <2/3, 2/3 to <1, fully exposed). This was calculated as the time in pregnancy spent using each type of stove rather than trimesters exposed. Similarly, in trial 2, not all pregnancies were exposed to each randomized intervention for the entire length of pregnancy. However, given that trial 2 was a parallel individually randomized trial, we analyzed the birth outcomes comparing the 2 randomized groups without regard to length of exposure in utero because these were equal in both groups. In trial 2, 55% were enrolled in the first trimester, 29% in the second trimester and 16% in the third trimester.

Intervention effects in trial 1 were estimated using linear regression to compare the mean differences and 95% confidence intervals (CIs) in birth weight and gestational age between levels of exposure to the improved stove using no exposure as the reference. Secular trends in birth outcomes were adjusted for, using a cubic spline with knots every 100 days. We used this approach because there is strong seasonality in birth weights and to some extent in gestational age in this population. For LBW, preterm, and SGA, relative risks and 95% CIs were calculated with no exposure as the reference, using Poisson regression with robust variance estimation. In trial 2, the same birth outcomes were compared between the 2 treatment groups (Envirofit and LPG stoves) using linear regression for continuous outcomes and Poisson regression for binary outcomes.
In trial 1, the study design randomized the timing of intervention rather than geographic clusters. Groups that received the intervention each month during stove installation were deliberately scattered geographically across the study area to reduce within cluster correlation. Therefore, we did not adjust for cluster randomization. In trial 2, the unit of randomization was the household. During this 1-year period, 5 sets of twins were born. Hence we do not believe adjusting for multiple births in the same household would impact the outcomes.

Both trials were approved by the institutional review boards of the Johns Hopkins Bloomberg School of Public Health and the Institute of Medicine, Tribhuvan University, Kathmandu, Nepal. Verbal informed consent was obtained from all participating households and individuals and documented on data collection forms. The trials are registered at ClinicalTrials.gov (NCT 00786877).

## RESULTS

A total of 2,379 live born infants were enrolled in trial 1 and 270 and 279 in the Envirofit and LPG stove groups in trial 2, respectively (Figures 1 and 2). Neither maternal nor infant characteristics differed between the randomized arms in trial 2 (Table 1). The only notable difference between

### TABLE 1. Baseline Characteristics of Mothers and Infants in 2 Randomized Trials of Improved Cookstoves in Rural Nepal

<table>
<thead>
<tr>
<th></th>
<th>Trial 1 (N=2379)</th>
<th>Vented Biomass (N=270)</th>
<th>LPG (N=279)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. (%)</td>
<td>No. (%)</td>
<td>No. (%)</td>
</tr>
<tr>
<td>Sex of newborn</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1216 (51.1)</td>
<td>149 (55.2)</td>
<td>149 (53.4)</td>
</tr>
<tr>
<td>Female</td>
<td>1163 (48.9)</td>
<td>121 (44.8)</td>
<td>130 (46.6)</td>
</tr>
<tr>
<td>Maternal age at delivery, mean (SD), y</td>
<td>2379 (23.5 (4.0))</td>
<td>270 (24.5 (3.7))</td>
<td>279 (23.0 (3.6))</td>
</tr>
<tr>
<td>&lt;18</td>
<td>157 (6.6)</td>
<td>15 (5.6)</td>
<td>17 (6.1)</td>
</tr>
<tr>
<td>18–35</td>
<td>2219 (93.3)</td>
<td>255 (94.4)</td>
<td>262 (93.9)</td>
</tr>
<tr>
<td>≥35</td>
<td>3 (0.1)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Parity, mean (SD)</td>
<td>2374 (1.7 (1.6))</td>
<td>266 (2.8 (1.7))</td>
<td>275 (1.6 (1.3))</td>
</tr>
<tr>
<td>0</td>
<td>578 (24.4)</td>
<td>40 (15.0)</td>
<td>53 (19.3)</td>
</tr>
<tr>
<td>1–3</td>
<td>1524 (64.2)</td>
<td>199 (74.8)</td>
<td>190 (69.1)</td>
</tr>
<tr>
<td>≥4</td>
<td>272 (11.5)</td>
<td>27 (10.2)</td>
<td>32 (11.6)</td>
</tr>
<tr>
<td>Maternal height, mean (SD), cm</td>
<td>2341 (150.1 (5.6))</td>
<td>245 (149.4 (5.4))</td>
<td>250 (150.3 (5.2))</td>
</tr>
<tr>
<td>&lt;145</td>
<td>357 (15.3)</td>
<td>42 (17.1)</td>
<td>29 (11.6)</td>
</tr>
<tr>
<td>145–149</td>
<td>654 (27.9)</td>
<td>81 (33.1)</td>
<td>83 (33.2)</td>
</tr>
<tr>
<td>150–154</td>
<td>800 (34.2)</td>
<td>80 (32.7)</td>
<td>91 (36.4)</td>
</tr>
<tr>
<td>≥155</td>
<td>530 (22.6)</td>
<td>42 (17.1)</td>
<td>47 (18.8)</td>
</tr>
<tr>
<td>Location of delivery</td>
<td>2353 (21.4)</td>
<td>161 (21.4)</td>
<td>169 (21.4)</td>
</tr>
<tr>
<td>Facility</td>
<td>483 (78.6)</td>
<td>95 (78.6)</td>
<td>97 (78.6)</td>
</tr>
<tr>
<td>Home or outdoors</td>
<td>1807 (1807)</td>
<td>95 (1807)</td>
<td>97 (1807)</td>
</tr>
<tr>
<td>Maternal education, mean (SD), y</td>
<td>2370 (1.5 (3.2))</td>
<td>270 (1.4 (3.2))</td>
<td>278 (1.9 (3.6))</td>
</tr>
<tr>
<td>0</td>
<td>1827 (77.1)</td>
<td>214 (79.3)</td>
<td>209 (75.2)</td>
</tr>
<tr>
<td>1–9</td>
<td>398 (16.8)</td>
<td>41 (15.2)</td>
<td>42 (15.1)</td>
</tr>
<tr>
<td>≥10</td>
<td>145 (6.1)</td>
<td>15 (5.6)</td>
<td>27 (9.7)</td>
</tr>
</tbody>
</table>

**Abbreviations:** LPG, liquid petroleum gas; SD, standard deviation.

*Some maternal ages at delivery were greater than the eligibility upper limit of 30 years at the time of enrollment.*
trials 1 and 2 was the percentage of deliveries at a facility, which was due to a strong secular trend in facility deliveries over the study time period. This trend was likely a result of a conditional cash transfer program implemented by the government of Nepal to incentivize women to deliver at facilities during this time.

In trial 1, the mean kitchen-based 1-day average PM$_{2.5}$ measurements in households using traditional stoves was 1380 (95% CI=1336, 1425) μg/m$^3$ (Table 2). The improved biomass stove reduced PM$_{2.5}$ to 936 (95% CI=895, 978) μg/m$^3$. In trial 2, the improved biomass stove had a similar PM$_{2.5}$ concentration to those in trial 1 at 885 (95% CI=810, 959) μg/m$^3$, while the LPG stoves had a lower mean concentration of 442 (95% CI=405, 482) μg/m$^3$. A small study done after the end of both trials to compare ambient kitchen to personal exposure in women of childbearing age showed that kitchen exposure overestimated personal exposure for traditional and improved biomass stoves, but was similar for LPG exposure (data not shown).

Adherence to the intervention was measured at weekly visits in both trials. In trial 1, at 90% of weekly visits, households reported using a stove other than the intervention stove (Envirofit) following its installation. In trial 2, study workers provided a standard canister of gas monthly. There were no gaps in delivery. In 11 of 53,007 weekly adherence visits in the LPG group, the household reported being out of gas. In 96 visits, the household reported that the gas regulator was damaged or not working. In trial 2, at 53% of weekly visits, households reported using a stove other than the LPG stove.

Trial 1 provided no evidence that any of the reproductive outcomes were affected by installation of the improved biomass stove (Table 3). Differences or relative risks were not statistically different from 0 or 1, respectively, and there was no dose response effect according to level of exposure. Even comparisons between no exposure in pregnancy versus exposure for the entire pregnancy did not indicate any effect on any birth outcomes. In trial 2, there was no statistically significant difference in mean birth weight, prevalence of LBW, preterm birth, or SGA among those pregnancies in the LPG households compared with Envirofit households (Table 4).

## DISCUSSION

Using a step-wedge design, we were unable to show an impact of an improved biomass stove on any of our measured birth outcomes. No differences were seen in a comparison of no exposure and full exposure to the intervention in pregnancy. The number of infants in each group was sizeable (588 [no exposure] versus 406 [full exposure] for birth weight within 72 hours, and 955 versus 574 for gestational age). Similarly, we

---

**TABLE 2. CO and PM$_{2.5}$ Kitchen-Based Concentrations in Households in 2 Randomized Trials Attributable to Improved Stoves, Adjusted for Potential Confounding, Rural Nepal**

<table>
<thead>
<tr>
<th>Measures of Exposure</th>
<th>Trial 1</th>
<th>Trial 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Traditional Biomass</td>
<td>Vented Biomass</td>
</tr>
<tr>
<td></td>
<td>N=2963</td>
<td>N=2752</td>
</tr>
<tr>
<td>Total daily PM$_{2.5}$, mean (95% CI), μg/m$^3$</td>
<td>1380 (1336, 1425)</td>
<td>936 (895, 978)</td>
</tr>
<tr>
<td>Time above 100 μg/m$^3$, mean (95% CI), hours, % sampling time</td>
<td>14.4 (14.2, 14.6), 66</td>
<td>11.7 (11.6, 11.8), 54</td>
</tr>
<tr>
<td>Time above 1000 μg/m$^3$, mean (95% CI), hours, % sampling time</td>
<td>5.3 (5.2, 5.4), 24</td>
<td>3.9 (3.8, 4.0), 18</td>
</tr>
<tr>
<td>Total CO, mean (95% CI), ppm</td>
<td>11.0 (10.6, 11.4)</td>
<td>6.7 (6.4, 7.1)</td>
</tr>
<tr>
<td>Time above 6 ppm, mean (95% CI), hours, % sampling time</td>
<td>6.6 (6.4, 6.8), 30</td>
<td>4.3 (4.1, 4.4), 20</td>
</tr>
<tr>
<td>Time above 9 ppm, mean (95% CI), hours, % of sampling time</td>
<td>4.9 (4.7, 5.0), 23</td>
<td>3.2 (3.1, 3.4), 15</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; CO, carbon monoxide; LPG, liquid petroleum gas; PM$_{2.5}$, particulate matter of 2.5 μm or less.

*Adjusted for rainfall and temperature on the day of measurement.*
were unable to show a difference between improved biomass and LPG stoves in a parallel randomized trial. For households with LPG stoves, the number of participants with exposure for their entire pregnancy and those with no exposure for their entire pregnancy was too small for any meaningful comparison in trial 2, but the sample size for the standard intent to treat was adequate to detect a meaningful difference.

In trial 1, the same women crossed over from one intervention to another and would therefore have had the same underlying risks and morbidities in the same pregnancy; hence we did not adjust for morbidity or other characteristics. Trial 2 was randomized to one intervention or the other. Table 1 shows comparability of demographic and maternal characteristics, and randomization would make it likely that other morbidities were comparable as well.

Meta-analyses of observational studies estimate about 30% to 80% increased risk of LBW and 80–100 g lower mean birth weight with high

---

**TABLE 3. Birth Outcomes, Adjusted**

<table>
<thead>
<tr>
<th>Time During Pregnancy Mother Exposed to New Stove, %</th>
<th>0</th>
<th>&lt;33</th>
<th>33–65</th>
<th>66–99</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of infants gestational age recorded</td>
<td>943</td>
<td>165</td>
<td>141</td>
<td>125</td>
<td>474</td>
</tr>
<tr>
<td>Gestational age, mean (SD), week</td>
<td>38.6 (2.7)</td>
<td>38.4 (3.1)</td>
<td>39.2 (2.0)</td>
<td>38.8 (2.7)</td>
<td>38.5 (2.7)</td>
</tr>
<tr>
<td>Gestational age, mean difference (95% CI), week</td>
<td>Reference</td>
<td>−0.51 (−1.03, 0.01)</td>
<td>0.27 (−0.30, 0.85)</td>
<td>−0.24 (−0.87, 0.39)</td>
<td>−0.75 (−1.36, −0.14)</td>
</tr>
<tr>
<td>Preterm, %</td>
<td>22.5</td>
<td>23.6</td>
<td>13.5</td>
<td>21.6</td>
<td>22.2</td>
</tr>
<tr>
<td>Preterm, RR (95% CI)</td>
<td>1.00</td>
<td>1.38 (0.97, 1.97)</td>
<td>0.81 (0.50, 1.32)</td>
<td>1.41 (0.91, 2.20)</td>
<td>1.66 (1.08, 2.57)</td>
</tr>
<tr>
<td>Number of infants with birth weight measured</td>
<td>588</td>
<td>133</td>
<td>116</td>
<td>104</td>
<td>360</td>
</tr>
<tr>
<td>Birth weight, mean (SD), g</td>
<td>2630 (443)</td>
<td>2628 (443)</td>
<td>2647 (418)</td>
<td>2676 (408)</td>
<td>2657 (439)</td>
</tr>
<tr>
<td>Birth weight, mean difference (95% CI), g</td>
<td>Reference</td>
<td>−12.8 (−107.1, 81.4)</td>
<td>−7.7 (−112.7, 97.4)</td>
<td>28.9 (−87.2, 145.0)</td>
<td>−5.5 (−122.6, 111.6)</td>
</tr>
<tr>
<td>Low birth weight, %</td>
<td>38.6</td>
<td>39.9</td>
<td>30.2</td>
<td>33.7</td>
<td>32.2</td>
</tr>
<tr>
<td>Low birth weight, RR (95% CI)</td>
<td>1.00</td>
<td>1.08 (0.82, 1.41)</td>
<td>0.83 (0.59, 1.17)</td>
<td>0.92 (0.63, 1.34)</td>
<td>0.92 (0.65, 1.30)</td>
</tr>
<tr>
<td>Number of infants with both gestational age and birth weight recorded</td>
<td>522</td>
<td>118</td>
<td>102</td>
<td>93</td>
<td>331</td>
</tr>
<tr>
<td>Gestational age, mean (SD), week</td>
<td>38.7 (2.6)</td>
<td>38.6 (25.6)</td>
<td>39.2 (2.0)</td>
<td>38.9 (2.8)</td>
<td>38.4 (2.7)</td>
</tr>
<tr>
<td>Gestational age, mean difference (95% CI), week</td>
<td>Reference</td>
<td>−0.39 (−1.01, 0.22)</td>
<td>0.16 (−0.53, 0.84)</td>
<td>−0.28 (−1.03, 0.47)</td>
<td>−0.97 (−1.73, −0.20)</td>
</tr>
<tr>
<td>Preterm, %</td>
<td>20.7</td>
<td>22.0</td>
<td>12.8</td>
<td>20.4</td>
<td>23.3</td>
</tr>
<tr>
<td>Preterm, RR (95% CI)</td>
<td>1.00</td>
<td>1.41 (0.91, 2.18)</td>
<td>0.84 (0.46, 1.53)</td>
<td>1.51 (0.86, 2.62)</td>
<td>2.01 (1.13, 3.56)</td>
</tr>
<tr>
<td>Small for gestational age, %</td>
<td>47.5</td>
<td>52.5</td>
<td>55.9</td>
<td>50.5</td>
<td>44.1</td>
</tr>
<tr>
<td>Small for gestational age, RR (95% CI)</td>
<td>1.00</td>
<td>1.14 (0.90, 1.44)</td>
<td>1.21 (0.95, 1.54)</td>
<td>1.11 (0.83, 1.48)</td>
<td>1.00 (0.74, 1.34)</td>
</tr>
</tbody>
</table>

**Abbreviations:** CI, confidence interval; RR, relative risk; SD, standard deviation.

a Adjusted for secular trend (cubic spline every 100 days) and sex of infant.

b Gestational age includes any measures 24 to <43 weeks (5.8% of gestational ages fell outside this range).

c Birth weight includes all measured within 72 hours.

d Gestational age and preterm among those with birth weights within 72 hours of birth.
exposure to indoor air pollution based on a mix of self-reported fuel use and measurements of PM$_{2.5}$ and CO. The randomized trial (Guatemala) found a 26% reduction in LBW after adjustment, but the sample size was small (a total of 179 infants weighed within 48 hours) and the authors used a per protocol analysis to account for variability in several factors. The trial in Nigeria found an 88 g higher mean birth weight in those exposed to clean versus biomass fuel. This difference was statistically significant, but the researchers did not report the impact of the intervention on LBW. Preterm prevalence was also lower in the clean fuel group, but the numbers were small (10 versus 16 cases). The indoor air pollution levels in Nigeria were much lower than those measured in either of the trials in Nepal. Our kitchen CO measures were higher for both traditional and improved biomass stoves than in Guatemala, but the measures in that study were personal rather than ambient. Hence it is difficult to compare exposures between Nepal and Guatemala. When comparing no exposure with full exposure to improved biomass stoves in trial 1, we found a 16% reduction in LBW and a 23 g increase in mean birth weight, neither of which were statistically different, and no dose response to exposure in utero.

Although some evidence exists for an effect of improved biomass stoves and theoretically clean fuel stoves such as LPG, the levels of exposure to the traditional stoves were exceedingly high in our study population relative to other studies. The improved biomass stove reduced ambient kitchen PM$_{2.5}$ and CO significantly, but the mean PM$_{2.5}$ was still more than 37 times higher than the WHO standard. LPG stoves reduced PM$_{2.5}$ exposure significantly from that of improved biomass stoves, but these levels were still 18 times higher than the WHO standard of 25 μg/m$^3$ or lower. Reasons for continued high exposures despite the use of clean fuels likely

### TABLE 4. Birth Outcomes by Randomization to Improved Stove Type, Trial 2, Rural Nepal

<table>
<thead>
<tr>
<th>Stove Type</th>
<th>Vented</th>
<th>LPG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of infants gestational age recorded</td>
<td>248</td>
<td>243</td>
</tr>
<tr>
<td>Gestational age, mean (SD), a week</td>
<td>39.2 (2.2)</td>
<td>39.0 (2.4)</td>
</tr>
<tr>
<td>Difference (95% CI)</td>
<td>-0.3 (-0.7, 0.2)</td>
<td></td>
</tr>
<tr>
<td>Preterm, a %</td>
<td>13.3</td>
<td>19.4</td>
</tr>
<tr>
<td>RR (95% CI)</td>
<td>1.45 (0.97, 2.19)</td>
<td></td>
</tr>
<tr>
<td>Number of infants with birth weight measured</td>
<td>188</td>
<td>207</td>
</tr>
<tr>
<td>Birth weight, mean (SD), b g</td>
<td>2780 (427)</td>
<td>2742 (431)</td>
</tr>
<tr>
<td>Difference (95% CI)</td>
<td>-37 (-122, 47)</td>
<td></td>
</tr>
<tr>
<td>Low Birth Weight, b %</td>
<td>23.4</td>
<td>31.4</td>
</tr>
<tr>
<td>RR (95% CI)</td>
<td>1.34 (0.97, 1.86)</td>
<td></td>
</tr>
<tr>
<td>Number of infants with both gestational age and birth weight recorded</td>
<td>176</td>
<td>184</td>
</tr>
<tr>
<td>Gestational age, mean (SD), c week</td>
<td>39.2 (2.4)</td>
<td>39.0 (2.3)</td>
</tr>
<tr>
<td>Difference (95% CI)</td>
<td>-0.2 (-0.7, 0.3)</td>
<td></td>
</tr>
<tr>
<td>Preterm, c %</td>
<td>15.3</td>
<td>19.6</td>
</tr>
<tr>
<td>RR (95% CI)</td>
<td>1.28 (0.81, 2.01)</td>
<td></td>
</tr>
<tr>
<td>Small for Gestational Age, c %</td>
<td>47.7</td>
<td>46.7</td>
</tr>
<tr>
<td>RR (95% CI)</td>
<td>0.98 (0.79, 1.21)</td>
<td></td>
</tr>
</tbody>
</table>
included concurrent use of an open burning biomass stove within the household, high levels of dust in the air, and contamination from adjacent households where traditional open-burning continued to be practiced. These results indicate the need to further study stove design and stove use behaviors, as well as how these behaviors influence the use of these stoves. Although we eliminated cost barriers to the use of LPG stoves (by providing both stove and gas supply), such barriers are real and would likely further compromise the effectiveness of this intervention.

The strengths of these trials include randomization to timing of intervention (trial 1) and standard randomization (trial 2). These 2 sequential trials provided different levels of exposure by which to assess health impacts, providing an ability to assess a dose response to the exposure.

Limitations
These trials have several limitations. First, gestational age was not measured using ultrasound but rather the date of the last menstrual period. This method could lead to some misclassification of preterm birth. A primary limitation of these trials was that the exposures, while lower than those due to the traditional stoves, were still very high with the improved biomass stoves. Although the LPG stoves further lowered exposure, these remained high. Use of a stove other than the intervention stove occurred at least once per week in about 90% of households in trial 1 and 50% in trial 2. Also, we did not measure ambient outdoor exposures, which could have contributed to the high exposures seen in both groups. Furthermore, many pregnancies were only partially exposed to the lower levels of PM_{2.5} and CO because the improved biomass or LPG stove was installed part way through the pregnancy. Sixty-seven percent of infants were weighed within 24 hours, 72% within 48 hours, and 75% within 72 hours. Because we did not measure weights at the time of birth, and because infants lose weight within a few days of birth, our estimates of LBW are likely overestimates of the true LBW. However, because of randomization, this bias should be similar in both groups. These constraints may explain the lack of effect on birth outcomes. Another limitation is that exposures were ambient kitchen rather than personal measures, and the kitchen measures were likely an overestimate of the personal exposures.

CONCLUSION
In summary, these analyses do not provide any evidence that the introduction of an improved biomass stove reduced adverse birth outcomes compared with open-burning biomass or LPG stoves. The reduction in PM_{2.5} and CO achieved by improved stoves was possibly inadequate to demonstrate any protective effect on this population. Despite clean fuel reducing indoor PM_{2.5} further, exposure levels were still very high, likely due to stove stacking and possibly high levels of ambient outdoor pollution, and may explain the lack of effect with this intervention. Further work is needed to assess the impact of lower exposures to PM_{2.5} and CO for the entire length of pregnancy on adverse birth outcomes.

Funding: Financial support was received from the National Institute of Environmental Health Sciences, United States National Institutes of Health (ES015538), the Thrasher Research Fund (D2830-4), the Global Alliance for Clean Cookstoves of the United Nations Foundation (UNF-12-380), and the National Institute of Child Health and Human Development (HD092411).

Competing interests: None declared.

REFERENCES


Peer Reviewed

Received: December 17, 2019; Accepted: May 26, 2020; First published online: July 17, 2020


© Katz et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are properly cited. To view a copy of the license, visit http://creativecommons.org/licenses/by/4.0/. When linking to this article, please use the following permanent link: https://doi.org/10.9745/GHSP-D-20-00011
Are We Using the Right Approach to Change Newborn Care Practices in the Community? Qualitative Evidence From Ethiopia and Northern Nigeria

Zelee Hill, Pauline Scheelbeek, Yashua Hamza, Yared Amare, Joanna Schellenberg

Key Findings

- In Ethiopia, high community-level exposure to consistent messages and the perceptions of community health workers (CHWs) drove behavior change.
- In Nigeria, exposure to messages was limited, CHW were less trusted, and behavior change was reported less frequently.

Key Implications

- The role of saturation and trust in behavior change in Ethiopia suggests that CHW counseling interventions should not be assumed to work through didactic provision of messages that focus on the behavior benefits.
- In settings with low coverage of CHW visits, low levels of message reinforcement from other sources, and low trust in CHWs, achieving behavior change may be difficult, even among people who receive visits.

ABSTRACT

Changing behaviors is usually a core component of the role of community health workers (CHWs), but little is known about the mechanisms through which they change behavior. We collected qualitative data from 8 sites in Ethiopia and northern Nigeria where CHWs were active to understand how they change newborn care behaviors. In each country, we conducted 12 narrative interviews and 12–13 in-depth interviews with recent mothers and 4 focus group discussions each with mothers, fathers, grandmothers, and CHWs. We identified 2 key mechanisms of behavior change. The first was linked to the frequency and consistency of hearing messages that led to a perception that change had occurred in community-wide behaviors, collective beliefs, and social expectations. The second was linked to trust in the CHW, obligation, and hierarchy. We found little evidence that constructs that often inform the design of counseling approaches, such as knowledge of causality and perceived risks and benefits, were mechanisms of change.

INTRODUCTION

Community health workers (CHWs) are central to efforts to reduce neonatal mortality in many low-income settings, and prenatal or postnatal home visits by CHWs are a key component of their role. A meta-analysis of 8 cluster-randomized studies of home visits by CHWs reported that such visits could reduce neonatal mortality by 11% (risk ratio 0.89; 95% confidence interval [CI]=0.85, 0.94), with the reduction being greater in settings with high mortality and low facility delivery rates. Although CHWs vary regarding payment and training, changing caregiver behaviors is usually a core component of their role.

In many cases, CHWs facilitate behavior change through the use of counseling cards that include written instructions for the CHW and pictures to help families understand and remember the benefit of the behavior. The mother is usually the target of the visits, but including other family members is often encouraged. The counseling approach often focuses on information-centric interpersonal communication, and it is informed by theoretical models developed in high-income settings. These theories are driven by constructs such as rationality, autonomy, self-interest, and material progression. This approach is exemplified in a recent review.
of caregiver behavior change in low- and middle-income countries that described the need for caregivers to know about the causal link between the behavior and survival, perceive that a risk exists, experience self-efficacy, have the resources to overcome environmental constraints, and decide to engage in the beneficial practice. It has been argued that in some low-income settings constructs such as interdependence, relationships, trust, obligation, and collective decision making may be more salient. There have been calls for more research to understand the mechanisms through which CHW interventions change behavior—the inner workings of the system—so they can be better designed and targeted.

Current evidence suggests that CHWs are much more than conduits of information. A common theme among studies exploring how CHWs effect change is that they can serve as brokers and connectors. They achieve this through connecting people with services, using local networks, engaging with influencers, and advocating for external resources or appealing to authorities. Having an empathetic, caring, and trusting relationship with the community is considered fundamental to the workers’ success. Trust comes from CHWs being embedded in the community—that is, an “insider”—whom community members identify with and have confidence in. Being an insider also means that CHWs understand and are direct witnesses to communities’ needs and challenges and are well placed to know how to address them. CHWs can engage with community members and build rapport with them; use appropriate terms to talk about health (e.g., through local idioms or songs); use their own personal testimonies of behavior change or act as role models; package benefits to be appealing and salient to community members; and identify relevant strategies for behavior change such as persistence and repetition. CHWs have also been seen to affect change by taking on out-of-scope tasks or using personal resources to facilitate their role.

CHWs also differ from other community members due to their training, skills, and association with the health system, which places them as “outsiders.” As such, they can have increased acceptance, respect, and status. Trust in CHWs as outsiders can be part of a broader system of hierarchical trust in authorities, which may result in community members feeling obliged to follow advice from CHWs and may reduce a community’s ability to make informed decisions.

Trust and credibility can be lost if CHWs do not meet community expectations, for example, if they lack supplies or demonstrate poor capabilities or when communities doubt their motives. Motives can be questioned when CHWs are perceived as not working for the interest of the community but as having alliances elsewhere, either because of the selection process or because there is a widescale distrust of authorities.

In this article, we report data from both CHWs and the community, exploring how CHWs change newborn care behaviors in Ethiopia and northeast Nigeria. We use an approach informed by realist evaluation to identify the mechanisms and the contextual factors that influence change. In Ethiopia, we focus on health extension workers (HEW) and the Health Development Army (HDA) leaders who were supported by the Last Ten Kilometers (L10K) program, and we compare our findings with those from northeast Nigeria around the role of community volunteers supported by the Society for Family Health (SFH).

**METHODS**

**Study Setting Selection and Characteristics**

Data were collected between March and June 2015 in the Southern Nations, Nationalities, and Peoples Region and the Amhara region in Ethiopia and in Gombe state in Nigeria. Table 1 shows the key characteristics of the CHWs in both sites. In Ethiopia, HEWs underwent 1 year of training, were paid about $100 a month, and served around 5,000 people. They provided health promotion and disease prevention and treatment, and they worked both at their health post and in the community. In Gombe, Nigeria, the CHWs were women from the Federation of Muslim Women’s Association in urban areas and traditional birth attendants in rural areas. They underwent 5–6 days of training and were volunteers, but they received incentives for accompanying women to facilities for delivery and for referring women with danger signs.

In both the Ethiopian and Nigerian sites, CHWs made prenatal and postnatal home visits, which included promoting neonatal care practices and service utilization. In Ethiopia, the HEWs also provided counseling during antenatal care at health posts, and all women were part of the HDA. The HDA members were meant to work in small groups with a leader whose role included helping families adopt the HEW messages by identifying pregnant women and linking them to services, holding monthly meetings, and running...
participatory learning and action cycles. The HDA leaders were selected by their group, were volunteers, and received an average of 15 days of training facilitated by the HEWs.31

With the assistance of L10K and SFH program staff, we identified 4 study sites in each country that had CHWs in place who were considered to be active. In Ethiopia, we selected 4 kebeles (districts) that had no unusual characteristics such as being near an industrial center or a tertiary hospital. In Nigeria, we selected 4 local government areas (LGAs) (districts)—2 urban and 2 rural. Insurgents were active at the time of data collection, which meant that the study team needed to return to the state capital each day, and LGA selection was limited to sites that were within a few hours drive of the state capital. In Ethiopia, data collectors were able to access study sites that were more remote. Table 2 shows the characteristics of the selected study sites. The access to health care described in the Table was determined by the fieldworkers based on distance to the nearest health centers, availability of transport, and the general road conditions.

Behaviors
Interventions that have the greatest impact on neonatal mortality are those that improve care in labor, during birth, in the first week of life, and for small and sick babies.35 Interventions to improve immediate newborn care such as thermal care (drying and wrapping, skin-to-skin care, delayed bathing), hygienic cord care, and early initiation of breastfeeding are central to these interventions.36 and CHWs often promote these behaviors. To explore behaviors in sufficient depth, we focused on thermal care and breastfeeding behaviors. Specifically, we focused on the following behaviors: drying the newborn and either placing it in skin-to-skin contact with the mother or wrapping the newborn after delivery if skin-to-skin contact was not possible; delayed bathing; and immediate and exclusive breastfeeding.37

Interview and Focus Group Discussion Guide Development
The interview guides included free-flowing questions, such as a description of labor, delivery, and newborn care, to allow unanticipated themes to emerge (Table 3). These questions were followed by theory-driven questions identified from the study theoretical framework (Figure), which hypothesizes a set of mechanisms through which CHWs could change neonatal breastfeeding and thermal care behaviors, and the contexts that may enable or constrain these mechanisms. The framework was developed using a realist evaluation approach,27 the COM-B (capability, opportunity, motivation, and behavior) model,38 a review of behavior change theories from a range of disciplines,39–42 a systematic review of neonatal care practices in sub-Saharan Africa,43 and discussions with implementers in the 2 countries. All guides were pretested and amended as needed.

Data Collection
Narrative and in-depth interviews (IDIs) were conducted with mothers at their houses.
group discussions (FGDs) were conducted with mothers, grandmothers, fathers, and CHWs in a neutral location, such as a school for community members and the health centers for CHWs. Having multiple respondent groups allowed us to capture different perspectives and to triangulate findings. We conducted narratives with mothers to understand their experiences and to see how events influenced each other and IDIs to capture perceptions. FGDs were conducted with 3–7 participants each to explore issues that would benefit from being discussed. We enhanced FGD interaction and reduced social desirability bias through activity-oriented exercises.44 Interviews and FGDs lasted between 45 minutes and 2.5 hours.

Community respondents were identified through community informants, such as women’s organization leaders, by the CHWs, at places of worship, and through snowball sampling. As we were interested in the mechanisms through which CHWs change behavior, we excluded respondents who had no contact with these cadres. In Ethiopia, due to the small numbers of HEWs in a kebele, all HEWs in the study kebele and those from neighboring kebeles were invited to participate in the FGDs. The SFH CHWs in Nigeria were identified with assistance from SFH and through snowball sampling. For mother interviews, we aimed to get a range of participants regarding age, educational level, parity, sex of newborn, and place of delivery.

Data were collected in local languages by 4–6 interviewers in each country. In Ethiopia, interpreters were occasionally needed to translate from Amharic into Silte. Interviewers were social science graduates with 2–17 years of qualitative research experience. They received 4 days of classroom training, which included a detailed review and discussion of the content of the guides, probing strategies, transcription, and ethics. This training was followed by a 2-day field pilot.

Interviewers approached potential community respondents in their homes and CHWs at their homes or health posts, explained the study, obtained written informed consent, and arranged a convenient time for the interview. Three respondents declined to be interviewed, stating that they were too busy. Interviews and FGDs were audio recorded, transcribed, and translated into English by the interviewers during data collection with conceptual and, where possible, semantic equivalence.45 Data collection stopped when additional interviews and FGDs did not provide new information. This point was determined by frequent transcript reviews and through discussions with interviewers. The sample size, respondent groups, and the interview content are shown in Table 2.

Ethical approval was granted by the Ministry of Science and Technology in Ethiopia, the National Research Ethics Committee and the Gombe State Government in Nigeria, and the London School of Hygiene & Tropical Medicine in the United Kingdom.

**Data Analysis**

Analysis began during data collection through regular team meetings and 2 data analysis
### TABLE 3. Data Collection Method, Sample Size per Country, and Content in Ethiopia and Nigeria

<table>
<thead>
<tr>
<th>Method</th>
<th>Sample</th>
<th>Content</th>
</tr>
</thead>
</table>
| Narrative interviews with recent mothers | 12/C15 | • Labor and delivery story  
• How was the newborn cared for, by whom, what influenced care, who made decisions  
• Perceived newborn care knowledge and skills compared with others  
• Contacts with, and advice from, health workers and CHWs, was advice new, did they agree with it, did it influence their behavior  
• Other advice received, agreement with the advice  
• Importance of what family and friends think of the care they give |
| In-depth interviews with recent mothers | 12–13/C15 | • Newborn care practices in the community and what influences these  
• Influence and importance of family and friends on care  
• CHW roles and their suitability for the role  
• Most significant newborn health changes in last 2 years |
| FGD with recent mothers | 4/C15 | • Pile sort of feeding and thermal care cards into practiced/not practiced and promoted/not promoted by CHW  
• Completion of a story of conflicting advice about delayed bathing  
• Most significant newborn health changes in last 2 years  
• Reaction to statements that CHWs work does not bring change, and that grandmothers are responsible for newborn care |
| FGD with grandmothers | 4/C15 | • Reaction to pictures of feeding and bathing practices  
• Role of grandmothers in newborn care and in decision making  
• Ranking of people who influence newborn care  
• Completion of a story of conflicting advice about delayed bathing  
• Most significant newborn health changes in last 2 years  
• Reaction to statements that grandmothers’ role is to support traditional practices, that CHWs know everything about newborn care and that mothers do not listen to grandmothers |
| FGDs with fathers | 4/C15 | • Reaction to pictures of feeding and bathing practices  
• Role of fathers in newborn care and in decision making  
• Completion of a story of conflicting advice about delayed bathing  
• CHW roles and their suitability for the role  
• Fathers’ knowledge of CHWs’ advice, and the advice they trust most  
• Most significant newborn health changes in last 2 years  
• Reaction to statements that mothers and fathers should decide on newborn care, that CHW visits involve fathers, that CHWs do not bring change, and that grandmothers’ role is to support traditional practices |
| FGD with CHWs | 4/C15 | • Pile sort of feeding and thermal care behaviors practiced/not practiced and important/not important by CHW  
• Most significant newborn health and work changes in the last 2 years  
• Successes and challenges in their work  
• Community reaction to them and their work  
• Reaction to statements that families are always happy to see the CHW, that CHWs do not bring change, and that families agree with delayed bathing advice |

Abbreviations: CHW, community health worker; FGD, focus group discussion.
workshops to discuss emerging themes, receive feedback on transcripts, remove or add questions, and increase reflexivity. These initial themes and the theoretical framework were used to develop a deductive coding template in NVivo. Interviews and FGDs were then coded inductively within these broad themes, with new themes added as they emerged. Coding and interpretation were done separately for each country by 4 of the senior researchers (ZH, YH, YA, and PS), who initially read through the transcripts to identify a first set of inductive themes and codes and to get an impression of the data as a whole. To enhance coding rigor and conceptual thinking, initial coding was done through consensus coding of the same transcripts, a code book was developed, and the team regularly met to discuss coding. Each analyst then took a set of transcripts to code. The study lead author (ZH) checked all coding using the coding stripe function of NVivo. Coding discussions focused around merging conceptually similar codes and examining patterns, links, and contradictions in the data. We triangulated findings between respondent groups and data collection methods to check credibility. Due to logistical constraints and literacy issues, we were unable to return transcripts to participants for comment.

**RESULTS**

Table 4 shows the characteristics of the narrative and IDI respondents. Respondents had a range of ages, education levels, parities, and religion. We did not achieve the planned diversity in place of delivery possibly because we were working in more accessible areas (close to the state capital in Gombe state and within walking distance of a motorable road in Ethiopia), in areas where CHWs were active, and with women who had received at least 1 visit by a CHW. In addition, CHWs assisted in identifying some respondents and may have favored those who delivered in a facility.

The FGD participant mothers varied in age (range 20–40 years), parity (range 1–8 children), education (none to secondary level), and ethnicity. The FGD participant fathers were older (range 28–61 years of age), and the FGD participant grandmothers were almost all uneducated.

Most families interviewed in Ethiopia reported adopting practices promoted by the HEWs and HDA leaders, which differed from past behaviors.

**Table 4**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Ethiopia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age range</td>
<td>20–40 yrs</td>
</tr>
<tr>
<td>Parity range</td>
<td>1–8 children</td>
</tr>
<tr>
<td>Education level</td>
<td>None to secondary level</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
</tr>
</tbody>
</table>

In previous time, the baby was bathed immediately after birth, but now, based on the teaching of health post and
Skin-to-skin care was the only thermal care practice that was reported as not being promoted, and this practice was new to most of the community respondents.

We identified 2 interlinked mid-range theories relating to how HEW interactions with families influenced the promoted practices in Ethiopia—saturation and trust theories. A third theory, facility behavior, was identified but applied only in the context of women who delivered in a facility.

The satisfaction theory describes a set of mechanisms that relate to families receiving consistent messages, reinforced by multiple sources (HEW, HDA leaders, and health workers) at multiple time points (before pregnancy, pregnancy, and at delivery) and through multiple channels (community meetings, home visits, antenatal care, and at delivery):

"We are told not to give anything until 6 months...[By] [HDA leader] [HEW], and we are told in the health center or when we go to the health post for vaccination [antenatal care]." —Mother, IDI, Kebele D, Ethiopia

"I heard this advice before, but when they give me additional advice it strengthen the idea...the HEWs advised me that the baby should not be bathed immediately after birth, and it was the same thing that the health professionals advised me." —Mother, IDI, Kebele A, Ethiopia

This saturation of messages led to high levels of community-wide familiarity with the promoted behaviors, with social diffusion further reinforcing the messages:

"I didn’t hear that with my ears but I heard a rumor that if you delivered at health facility you have to bathe [the baby] after 3 days." —Mother, Narrative I, Kebele D, Ethiopia

This familiarity led to a perception that there had been a change in collective beliefs and practices and a social expectation that the new behaviors should be practiced:

"No one [in the family or community] suggested bathing the baby immediately ... They all know that the baby should be bathed after spending a day." —Mother, Narrative, Kebele C, Ethiopia

"It is the same [people’s behavior], it has no difference. They [others] have received the lesson...They have received the same lesson as me." —Mother, Narrative, Kebele C, Ethiopia

The HEWs attributed message penetration to the formation of the HDA, whose members were able to have more frequent contacts with households. This attribution was confirmed by some community respondents.

Having messages provided from multiple sources meant that even if a worker was less active, messages were still heard. For example, in 1 study site, HEWs were reported as being relatively inactive in making home visits, but most women still received messages from other sources.

Grandmothers were also familiar with the promoted behaviors and, with the exception of delayed bathing, were generally supportive of them, recognizing that this was a new time:
It is our children who advise us about what they know . . . Today’s mothers are wise. In previous time, there was no education. —Grandmother, FGD, Kebele D, Ethiopia

As a result of being “educated” about the practices, mothers had increased maternal agency and felt more knowledgeable than their “outdated” elders and more able to reject their advice:

These days, they [HEW] give us good education on how to take care of babies . . . my families are getting older . . . So even though they give us advice, we would not follow that . . . I follow my own way. —Mother, Narrative, Kebele B, Ethiopia

Overall, fathers were less familiar with the details of the behaviors being promoted, but they were supportive of the CHWs and health worker advice, as illustrated in this conversation in one FGD:

Previously . . . our mothers decided. —Father A, FGD, Kebele A, Ethiopia

Now I can decide that it [food] should not be given before 6 months. —Father B, FGD, Kebele A, Ethiopia

Because health professionals teach this. —Father A, FGD, Kebele A, Ethiopia

Familiarity with the messages was important, but knowledge of the benefit of the behavior was not a prerequisite for adoption. For example, although many respondents could report that delayed bathing helped maintain warmth, several did not know the reason for the practice but still reported that they carried it out. In these cases, the trust theory was the driver of change:

I do what they told me to do; I waited the time and bathed the baby . . . I was thinking it could not be harmful; they said this because it is useful. —Mother, IDI, Kebele D, Ethiopia

The trust theory relates to a trust in the CHWs and health workers as a collective to act in the community’s interest, leading families to follow the advice given:

They [mothers] listen to the HEW, even if they tell you that you have to walk naked. —Grandmother, FGD, Kebele C, Ethiopia

Trust was related to CHWs’ high status in that they were “better” and “more educated” than community members:

I was advised about it before delivery. I accept what they told me because they know better. —Mother, Narrative, Kebele C, Ethiopia

We just respect educated people. They are educated, and we are just farmers. —Mother, IDI, Kebele A, Ethiopia

Trust was also related to CHWs being seen as connected to the government, as well as to their visibility, connectedness, and understanding of the community.

Although the HEWs as a collective were trusted as acting in the community’s interest, this trust could be lost for individual HEWs, for example, if they provided inequitable services, had a poor work ethic, took bribes, or were unkind:

They [HEWs] are not responsible enough to carry out their tasks . . . they want you to give them something as a bribe . . . if she doesn’t give something as a bribe, the woman won’t get any service . . . they did not have good manner. —Mother, IDI 5, Kebele B, Ethiopia

Trust in the HEWs and their status were combined with a hierarchical relationship that resulted in families feeling obligated to follow the advice of HEWs. This circumstance is illustrated by community respondents describing behaviors as “forbidden” or “not allowed” and reporting that the HEWs “commanded” them. This hierarchical relationship was compounded by a belief that the HEWs had the power to make life difficult if advice was not followed, either by being angry with families or denying them services.

HDA leaders did not have the same status as HEW/health workers and were described as “villagers just like us.” They were viewed as conduits of information and were important drivers of the saturation mechanisms but were not drivers of the trust mechanisms:

They give advice as local people, not different from advising as a neighbor. —Father, FGD, Kebele B, Ethiopia

The third theory, centered on facility behavior, relates to the recent increase in facility deliveries, which meant that some family practices had become health worker practices:

When they deliver at the health center, they practice this [immediate breastfeeding] . . . they practice it mostly because they deliver at the health facility. —CHW, FGD, Kebele A Ethiopia

At a facility delivery, immediate drying and wrapping were done by the health workers, who also could insist on early breastfeeding and no bathing during the stay in the facility.

The 3 theories described above were not triggered equally for all behaviors, with delayed
bathing being the most difficult behavior to change. This was related to some families, and in particular grandmothers, having strong culturally entrenched beliefs against leaving the baby “dirty” after birth. This was especially the case if the baby was bloody or had an obvious vernix:

We will bathe the baby anyway, how can we sit idle without bathing the baby. —Mother, IDI, Kebele D, Ethiopia

Another contextual factor that enhanced the uptake of behaviors was the value given to education and modernity:

Do you think I want to do like in the past? I want it to be just like the modern time . . . Today’s mothers are young and modern. They easily accept new ideas. —Mother, Narrative Kebele C, Ethiopia

In the Nigerian sites, there were fewer reports of behaviors being changed, and the impact of saturation was less striking than in Ethiopia, as messages were received less often and from fewer sources. There was less evidence of a change in collective beliefs as well as more cultural barriers to behavior change, especially when mothers received conflicting advice:

. . . I gave water on the third day . . . I am confused with the advice I am getting, some say give, some say do not give. —Mother, Narrative, LGA C, Nigeria

As in Ethiopia, CHWs were generally trusted in the Nigeria sites, but fewer respondents reported that they carried out behaviors without knowing their benefits:

I’ll follow the ways that I already know, I’ll follow the old ways until I hear more explanation on it. —Mother, Narrative, LGA C, Nigeria

As in Ethiopia being “educated” by a CHW allowed some mothers to challenge the advice of elders, and a theme in the grandmother FGDs emerged around times changing and their advice no longer being relevant:

We the old grandmothers are in trouble. We have been swept into the dustbin. —Grandmother, FGD, LGA A, Nigeria

But in many households, elders remained key influencers of newborn care with a strong family hierarchy. For example, several mothers reported they followed or pretended to follow the advice of the elders to ensure harmonious relationships.

Unlike in Ethiopia, the facility behavior theory was not a key mechanism of change in Nigeria, as health staff did not practice the behaviors that CHWs were promoting as frequently as in Ethiopia:

At hospital when you deliver, you will be allowed to rest for 1 hour . . . which in most cases, you will be sleeping. So, you see you don’t even have the time to breastfeed your baby. —Mother, IDI, LGA D, Nigeria

However, a theme emerged around a facility delivery giving women more autonomy in decision making because they were viewed as having adopted and being exposed to modern practice:

She [grandmother] used to say I should breastfeed them, and about water . . . I should give them . . . I tell her that in the hospital they told us not to give and she will not force me, she will respect what the hospital says. —Mother, Narrative, LGA A, Nigeria

As in Ethiopia, the mechanisms were not triggered for all behaviors in Nigeria, but in contrast to Ethiopia, delayed bathing was more amenable to behavior change than breastfeeding practices. This was because families had been convinced that the baby could be cleaned adequately by rubbing them with oil rather than bathing them, making delayed bathing culturally acceptable:

They [health worker] clean the baby very well, so baby will not smell, if they leave the baby dirty . . . then, of course the baby will smell, my baby was very clean so, no problem. —Mother, Narrative, LGA B, Nigeria

### DISCUSSION

The saturation and trust theories that we identified in our study sites are similar to findings from previous studies on how CHWs affect change, but are in contrast to the idea that knowledge of a causal link and perception of a risk are needed for behavior change. Many behavior change theories include constructs related to psychosocial, environmental, and social network determinants of behaviors, yet there can be an overreliance on didactic communication in health promotion.

The saturation mechanism was linked to frequent and consistent information, which was provided more often in Ethiopia than in Nigeria. Where it occurred, saturation led to a perception that there had been a change in community-wide behaviors, collective beliefs, and social expectations of behaviors. CHW counseling interventions usually focus on changing the behavior of one individual at a time, with less emphasis given to

In the Nigerian sites, there were fewer reports of behaviors being changed, and the impact of saturation was less striking than in Ethiopia.

In many households, elders remained key influencers of newborn care, with a strong family hierarchy.
Inadequate training and support, high workloads, and poor motivation and performance can influence CHWs’ ability to deliver behavior change interventions.

Message saturation or exposure. CHWs themselves have identified the need for persistence and repetition, and when sufficient in number, positive testimonies from community members about CHWs’ work have been found to create a snowball effect and enhance community-wide trust.

In the Ethiopian sites, high exposure to consistent information was achieved, but obtaining the same outcome may be difficult in many settings. Few countries have reached more than 10% coverage for postnatal home visits, and some evidence suggests that in some settings, CHWs lacked the time, motivation, and support to engage with high-quality behavior change. Behavior change may be difficult to achieve in settings with less functional CHW programs, which may in some part explain the difference in the findings between Ethiopia and Nigeria. This difference highlights the importance of context in the design of counseling interventions within CHW programs.

The trust mechanism was related to trust in CHWs, status, obligation, and hierarchy but also to CHWs’ closeness to and understanding of the communities they serve. In Ethiopia, this mechanism was seemingly more important than individuals being convinced of the specific benefit of a behavior. Previous studies have highlighted the importance of relationships, trust, and power between CHWs and the community, and the impact that these factors can have on acceptability, motivation, and performance. We demonstrate the centrality of this relationship to behavior change, and a greater emphasis on it may be warranted in programs. The trust mechanism was likely also influenced by the unique Ethiopian social and political context at the time of the study, which was characterized by a stratified and hierarchical society with strong political control and high levels of social cohesion. The role of hierarchical trust may be particularly strong in Ethiopia compared with Nigeria and other settings, which may have resulted in caregivers not making informed decisions but feeling obligated to comply with the CHWs advice.

In the Nigerian sites, respondents reported that messages were received less often and from fewer sources. In addition, the advice of CHWs was questioned more, and elders remained as key influencers. These findings could be attributed to a poorer functioning CHW system in Gombe, with lower coverage than in Ethiopia, and the lower training and the volunteer nature of the CHWs, as well as the broader cultural, political, and health systems contexts. For example, in Ethiopia, the HEWs provided services both in the community and at the health post, which may have enhanced their status and their ability to influence the community compared with the CHWs in Nigeria who only had a community role. In addition, consistent messaging may have been harder to achieve in Nigeria, which was characterized by interventions with disjointed designs, a proliferation of strategies, and multiple health systems, which created unique challenges. Further, there may have also been less hierarchical trust in authorities.

Limitations
We used multiple study sites, purposive sampling to saturation, reflexivity, triangulation of methods and respondent groups, consensus coding, and within and across case analysis to increase data quality and the transferability of findings. However, the findings may not be transferrable to settings with very different contexts, especially those that are more remote and inaccessible. This limitation is particularly the case in Nigeria, where our site selection was restricted by insurgent activities. The inclusion of more women who delivered at home would have allowed us to examine differences in mechanisms by place of delivery. The potential exists that reporting was influenced by social desirability bias, a particular problem in Ethiopia. In addition, the respondents identified by CHWs may have had different attitudes and experiences compared with families that were perhaps less favored by the CHWs.

Conclusion
Our study did not aim to generalize about the effectiveness of the CHW programs but to uncover the mechanisms through which they may work when they are functioning. CHW programs face issues with inadequate training and support, high workloads, and poor motivation and performance, and these issues have been found to influence CHWs’ ability to deliver behavior change interventions. The mechanisms we uncovered are unlikely to be triggered unless health systems function adequately and coverage and quality are improved.
Funding: The research was funded by the Bill & Melinda Gates Foundation as part of the IDEAS (Informed Decisions for Actions to improve maternal and newborn health) study.

Competing interests: None declared.

REFERENCES


How Community Health Workers Change Behaviors


Global Health: Science and Practice 2020 | Volume 8 | Number 3
The Critical Role and Evaluation of Community Mobilizers in Polio Eradication in Remote Settings in Africa and Asia

Judy Lewis, Karen LeBan, Roma Solomon, Filimona Bisrat, Samuel Usman, Ahmed Arale

Key Findings

- Data use, community engagement, local adaptation, linkage with the health system, and a strong community platform are critical for successful community programming.
- Community-based disease surveillance using local volunteers enhanced national and district efforts.

Key Implications

- Program managers should consider using the Updated Program Functionality Matrix for Optimizing Community Health Programs of the Community Health Worker Assessment and Improvement Matrix (AIM) tool to develop, implement, and assess community health worker programs.
- Policy makers should support and strengthen strong community platforms in disease eradication programs that engage communities, promote local ownership, and use community-level workers.
- Emerging infectious diseases will require community engagement through local mobilizers to implement successful government prevention and response efforts.

ABSTRACT

This article assesses the CORE Group Polio Project (CGPP) experience over a 20-year period in 5 countries. It examines how a program designed to provide social mobilization to eradicate one disease, and which did so effectively, functioned within the general framework of community health workers (CHWs). Vertical health programs often have limited impact on broader community health. CGPP has a 20-year history of social mobilization and effective program interventions. This history provided an opportunity to assess how CGPP community mobilizers (CMs) functioned in polio and maternal and child health. The Updated Program Functionality Matrix for Optimizing Community Health Programs tool of the CHW Assessment and Improvement Matrix (AIM) was used to examine CGPP CM roles across different contexts. The analysis determined that CGPP CMs met the basic level of functioning (level 3) for 6 of the 10 components of the AIM tool. This cross-country descriptive analysis of the CGPP demonstrates the importance of embracing the full range of CHW AIM components, even in a vertical program. Use of data, community involvement, local adaptation, and linkage with the health system are especially critical for success. This general lesson could be applied to other community mobilization and disease/epidemic control initiatives, especially as we face the issues of the COVID-19 pandemic.

INTRODUCTION

This article examines how a program designed to provide social mobilization to eradicate polio, and which did so effectively, functioned within the general framework of community health workers (CHWs).

Although vertical health programs often have limited impact on broader community health, we wanted to assess how well the CORE Group Polio Project (CGPP) community workers functioned in the areas of polio and maternal and child health. We also wanted to examine their roles in different contexts using the components of the Updated Community Health Worker Assessment and Improvement Matrix (CHW AIM).

CGPP has a 20-year history of social mobilization and effective program interventions. Published external evaluations and peer-reviewed articles about CGPP have demonstrated substantial success in increasing oral polio vaccine (OPV) 0 (the newborn dose) and routine OPV...
and immunization coverage as well as detecting acute flaccid paralysis (AFP) in hard to reach and resistant populations (Table 1).9,10 The experience of CGPP offers many lessons about implementing vertical programs, developing and deploying a cadre of community-level workers, and engaging with the health care system.

SITUATION OF POLIO AT THE INCEPTION OF CGPP
Mass immunization campaigns in the mid-1990s achieved high levels of polio immunization coverage. However, in some countries, there remained important pockets of children who were repeatedly missed and served as residual pockets of continuing transmission. By 1999, polio geography and incidence had decreased considerably but it was clear that the goal of eradicating polio by 2000 would not be met and that more focused efforts would be needed to address polio “hotspots” (see Losey et al.11).

Where community involvement was low, OPV coverage remained low. Although millions of temporary volunteers supported mass campaigns, their job ended when the campaign ended. Conflict, political instability, geographic inaccessibility, nomadic and mobile populations, poor infrastructure, and anti-vaccination social and religious beliefs were some of the obstacles that led communities either to refuse immunization or prevent participation, resulting in low routine immunization and OPV rates. Polio experts within United States Agency for International Development (USAID) began to make recommendations for broad social mobilization efforts to increase community participation in the eradication of polio. Experts increasingly recognized that each remaining polio-endemic country offered a unique set of challenges that required local solutions.12 CGPP started in 1999 with funding from USAID to address these issues. This process has been well described.11

METHODOLOGY
We used a mixed-methods evaluation approach to compare the evolution of CHWs within CGPP. The 2 first authors collected the data. We began with an extensive literature and document review about CGPP, which included multiple mid- and final evaluations reflecting different grant periods. This review was a major undertaking because of the project’s 20-year history in 11 countries. We also conducted surveys through computer, phone, and in-person interviews with CGPP secretariat directors and staff. The survey included 18 broad questions about program operation, management, and development over time, focusing on community mobilizer (CM) roles. The timeframe for data

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Angola</th>
<th>Ethiopia</th>
<th>India</th>
<th>Kenya</th>
<th>Somalia</th>
<th>Nigeria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-polio acute flaccid paralysis rate per 100,000 children under age 15 within 14 days of onset of paralysis with 80% or better stool adequacy</td>
<td>Not available</td>
<td>2.2 (2012) to 2.8 (2017) exceeding national rate of 2.5 (2017)7</td>
<td>Not applicable</td>
<td>2.5 (2017)4</td>
<td>4 (2017)4</td>
<td>13.6 (2014) to 19.6 (2017)8</td>
</tr>
</tbody>
</table>

Abbreviation: OPV, oral polio vaccine; OPV0, oral polio vaccine newborn dose; OPV3, oral polio vaccine third dose.
The focus was on 5 programs: India, Ethiopia, and Angola, which have been in operation the longest, and Nigeria and Kenya/Somalia, where conflict and migration were challenges.

The lead authors were involved in all levels of data collection and analysis. We focused on 5 programs: India, Ethiopia, and Angola, which have been in operation the longest, and Nigeria and Kenya/Somalia (Horn of Africa Program), which are more recent but face particularly challenging situations.

We used the Updated Program Functionality Matrix for Optimizing Community Health Programs\(^\text{13}\) of the CHW AIM\(^\text{14}\) as the framework for our analysis. CHW AIM uses 10 programmatic components that have been found to contribute to an effective CHW program. Each of the 10 components is subdivided into 4 levels of functionality: (1) nonfunctional, (2) partially functional, (3) functional, and (4) highly functional. The program also includes a process for creating a participatory functionality score, which we did not use. We used the criteria for level 3 (functional) to examine whether the CM work met the criteria for this level of functioning. This standard was used across country programs and contexts for each of the 10 components to examine the long-term impact of CM roles in each country.

**CGPP COUNTRY PROGRAMS**

The 5 CGPP country programs discussed in this article began between 1999 and 2014. Program inception dates and the number of collaborating nongovernmental organizations (NGOs) are provided in Table 2.

**Name and Number of Community Mobilizers**

CGPP’s CHWs were CMs, which are sometimes referred to by different names. The term CM is used in this article for all CGPP country programs to distinguish the CGPP cadre from other CHWs used by NGOs and government agencies. Social mobilization was the main strategy to provide polio education, engage communities in polio vaccination, track children missed during OPV campaigns, and conduct AFP surveillance in high-risk populations.

The country-specific names used for the CMs reflected government policy or nomenclature widely used by partners when the program began (Table 3).

As governments developed community health strategies, the project incorporated government CHWs into their programs to help address polio.

---

**TABLE 2.** CORE Group Polio Project Country Start Dates and Number of Collaborating NGOs, Past and Present\(^a\)

<table>
<thead>
<tr>
<th>Country</th>
<th>Year started</th>
<th>Number of international NGOs</th>
<th>Number of local NGOs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angola</td>
<td>1999</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>2001</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>India</td>
<td>1999</td>
<td>6</td>
<td>77</td>
</tr>
<tr>
<td>Kenya/Somalia</td>
<td>2014</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Nigeria</td>
<td>2013</td>
<td>3</td>
<td>8</td>
</tr>
</tbody>
</table>

Abbreviation: NGO, nongovernmental organization.

\(^a\) The NGOs did fluctuate over time and area covered, so for all data in this article, we have referenced numbers from Losey et al.\(^\text{11}\)

**TABLE 3.** Name, Number, and Type of Community Mobilizers by CORE Group Polio Project Country

<table>
<thead>
<tr>
<th>Country</th>
<th>Name</th>
<th>Current Number</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angola</td>
<td>Community volunteers</td>
<td>2,700 (2017 FE)</td>
<td>Part time</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>Community volunteers</td>
<td>13,720 (2017 FE)</td>
<td>Part time</td>
</tr>
<tr>
<td>India</td>
<td>Community mobilization coordinators</td>
<td>1,100 (2017 FE)</td>
<td>Part time</td>
</tr>
<tr>
<td>Kenya/Somalia</td>
<td>Community health volunteers</td>
<td>1,025 (2017 FE)</td>
<td>Part time</td>
</tr>
<tr>
<td>Nigeria</td>
<td>Volunteer community mobilizers</td>
<td>2,200 (2017 FE)</td>
<td>Part time</td>
</tr>
</tbody>
</table>

Abbreviation: FE, Final Evaluation.
For example, as the Ethiopia government deployed its Women’s Development Army (WDA), CGPP worked with the volunteer WDA leaders (1 for every 30 WDA volunteers) in CGPP implementation areas where WDA volunteers were active, and the number of CMs greatly increased. Numbers of CMs varied over time as partners, population, or geographic area changed. For example, in Ethiopia, CGPP trained 2,000 CMs between 2004 and 2006; 4,165 between 2007 and 2012; and 13,720 between 2013 and 2017 (this included WDA leaders beginning in 2015).

All projects had part-time CMs who fit the description for CHW-regular described by Hodgins et al.15

Location of Work and Population Reached
In each country, CGPP worked in areas assigned by the in-country Interagency Coordinating Committee (ICC) for Polio Eradication (Table 4).

Areas and population reached with OPV changed during the project, often on short notice, based on reviews of immunization data and/or the need to reach special at-risk and inaccessible populations. Populations were large and often in noncontiguous areas. For example, over 1.4 million people were reached in Nigeria through social mobilization efforts.4

### ANALYSIS OF CM FUNCTIONALITY IN THE CHW AIM

The following sections highlight key programmatic components of CM functionality as described in the Updated CHW AIM. Table 5 provides a summary of the matrix and the criteria used to determine whether AIM level 3 functionality was achieved for each component. Similarities and differences between programs are provided below.

#### Role and Recruitment

**Recruitment.** Table 6 shows that in all countries, CGPP staff provided generic selection criteria to NGOs and/or community leaders, and communities played an important role in identifying candidates. Communities could modify the criteria to best match the local context, such as literacy or sex of the CM. In terms of who chose the CMs, NGOs selected the CMs in Angola since they were already working with them. In Kenya/Somalia, health facility staff also participated in the selection process. In Angola, India, and Kenya/Somalia, the NGO made the final hiring decision. In Ethiopia, community leaders made the decisions with input from health facility staff and district administrators (later by the health extension workers). In Nigeria, ward selection committees decided. In terms of the types of people selected, existing CMs or CHWs were selected when possible in Angola, Kenya/Somalia, and India. In Ethiopia, community leaders and influencers (often religious figures) were selected. Madrasa teachers and elected officials initially helped identify candidates in India to recruit Muslim women.

**Selection Criteria.** Table 7 identifies defining characteristics of CMs. These characteristics varied by cultural context and changed over time. Initially, literacy was prioritized, but later characteristics such as respect and trust by the communities were emphasized.

---

### TABLE 4. Location of CORE Group Polio Project Work and Population Reached (Annual Reports and 2017 Final Evaluations)

<table>
<thead>
<tr>
<th>Country</th>
<th>Location</th>
<th>Population Reached</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angola</td>
<td>5 provinces (Benguela, Bie, Cuanza Sul, Lunda, Malange)</td>
<td>&gt;9 million children under 15</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>85 districts in 5 regions (Benshangul-Gumuz; Gambella; Oromiya; Southern Nations, Nationalities and Peoples; Somali) 185 border crossing pointsa</td>
<td>&gt;6 million people of which 1,806,950 are children under age 5a</td>
</tr>
<tr>
<td>India</td>
<td>58 blocks in 12 high-risk districts of Uttar Pradesh, 2 districts in Assam, and 1 district in Haryana</td>
<td>600,000 households reaching population of 3 milliona</td>
</tr>
<tr>
<td>Kenya/Somalia</td>
<td>Kenya: 7 counties (Lamu, Garissa, Mandera, Marsabit, Turkana, Wajir, and parts of Nairobi) Somalia: 3 border regions (Lower Juba, Gedo, and Bakool)</td>
<td>Kenya: 466,250 children under age 5 Somalia: 109,000 children under age 5</td>
</tr>
<tr>
<td>Nigeria</td>
<td>32 local government areas in 5 northern states (Borno, Kaduna, Kano, Katsina, and Yobe) 6 internally displaced persons camps</td>
<td>Approximately 500,000 children under age 5a</td>
</tr>
</tbody>
</table>

---

a Data from secretariat directors.
TABLE 5. Community Health Worker Assessment and Improvement Matrix Tool Components and Criteria Used for CORE Group Polio Project

<table>
<thead>
<tr>
<th>CHW AIM 2018: Revised Programmatic Components</th>
<th>CHW AIM 2018 Elements Examined for CGPP CMs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Role and Recruitment</td>
<td>Recruitment:</td>
</tr>
<tr>
<td>How the community, CHW, and health system design and achieve clarity on the CHW role and from where the CHW is identified and selected.</td>
<td>• Initial selection</td>
</tr>
<tr>
<td>Level 3 requires:</td>
<td>• Final decision</td>
</tr>
<tr>
<td>Recruitment: CHWs recruited from community and community consulted in selection. Criteria for functionality, attitudes, expertise, and availability of CHWs clearly delineated.</td>
<td>• Type of CM</td>
</tr>
<tr>
<td>Role: Clearly defined and documented, agreed upon by CHW, community, and health system.</td>
<td>• Community mobilization to increase polio and routine vaccination rates</td>
</tr>
<tr>
<td>Workload and location: CHW to population ratio reflects expectations, population density, geographical constraints, and travel requirements.</td>
<td>• Community-based surveillance of acute flaccid paralysis</td>
</tr>
<tr>
<td>2. Training</td>
<td>Continuing education:</td>
</tr>
<tr>
<td>How preservice training is provided to CHWs to prepare for their roles and to ensure they have the necessary skills to provide safe and quality care. How ongoing training is provided to reinforce initial training, teach CHWs new skills, and help ensure quality.</td>
<td>• Trainers</td>
</tr>
<tr>
<td>Level 3 requires:</td>
<td>• Content of training</td>
</tr>
<tr>
<td>Initial training: meets global guidelines and occurs within 6 months of recruitment.</td>
<td></td>
</tr>
<tr>
<td>Continuing education: provided at least annually and vertical topics are integrated</td>
<td></td>
</tr>
<tr>
<td>3. Accreditation</td>
<td>Assessment of CM health knowledge and competencies</td>
</tr>
<tr>
<td>How health knowledge and competencies are assessed and certified prior to practicing and recertified at regular intervals while practicing.</td>
<td>• External program evaluations</td>
</tr>
<tr>
<td>Level 3 requires:</td>
<td></td>
</tr>
<tr>
<td>CHW health knowledge and competencies are tested and a minimum standard must be met.</td>
<td></td>
</tr>
<tr>
<td>4. Equipment and Supplies</td>
<td>Continuous supply of job aids</td>
</tr>
<tr>
<td>How the requisite equipment and supplies are made available when needed to deliver expected services.</td>
<td></td>
</tr>
<tr>
<td>Level 3 requires:</td>
<td></td>
</tr>
<tr>
<td>Equipment, supplies, and job aids are provided and available for resupply on a regular basis.</td>
<td></td>
</tr>
<tr>
<td>5. Supervision</td>
<td>Type of supervisor</td>
</tr>
<tr>
<td>How supportive supervision is carried out such that regular skill development, problem solving, performance review, and data auditing are provided.</td>
<td>• Average number CMs supervised</td>
</tr>
<tr>
<td>Level 3 requires:</td>
<td>• Supervisor paid</td>
</tr>
<tr>
<td>A dedicated trained supervisor uses checklists to conduct supervision visits at least every 3 months and uses summary statistics to identify areas for improved service delivery.</td>
<td>• Tools used</td>
</tr>
<tr>
<td>6. Incentives</td>
<td>Frequency of supervision performance evaluation (individual and program)</td>
</tr>
<tr>
<td>How a balanced incentive package reflecting job expectations, including financial compensation in the form of a salary and nonfinancial incentives, is provided.</td>
<td></td>
</tr>
<tr>
<td>Level 3 requires:</td>
<td>Financial (honorarium, transport/food allowance)</td>
</tr>
<tr>
<td>CHWs are compensated at a competitive rate and receive nonfinancial incentives</td>
<td>• Nonfinancial (certificates, performance awards, formal recognition, skill development, uniforms, job aids, free access to health services)</td>
</tr>
<tr>
<td></td>
<td>• Community recognition</td>
</tr>
</tbody>
</table>

Global Health: Science and Practice 2020 | Volume 8 | Number 3
community and knowledge of local customs and norms were found to be more important by CGPP. NGO staff developed methods for working with CMs with low literacy, such as having their children write in immunization registers or using oral storytelling with the supervisors.

India preferred female community health volunteers since they have easier access to mothers, the principal caregivers for children. In Ethiopia, there were initially more women when the project worked in agrarian districts. When the project moved to border areas and pastoralist and semipastoralist districts, the ratio shifted to equal numbers of women and men. Reasons included religious preference, security issues, women being too busy with household work, and community decision making. The ratio changed again as the government required all WDA volunteers to be women. In Nigeria, project surveys validated that women were seen as the primary caregivers.

**TABLE 5. Continued**

<table>
<thead>
<tr>
<th>CHW AIM 2018: Revised Programmatic Components</th>
<th>CHW AIM 2018 Elements Examined for CGPP CMs</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Community Involvement</td>
<td>Discuss CM role and selection</td>
</tr>
<tr>
<td>How a community supports the creation and maintenance of the CHW program.</td>
<td>Provide feedback on performance</td>
</tr>
<tr>
<td>Level 3 requires:</td>
<td>Solving problems</td>
</tr>
<tr>
<td>Community supports, recognizes, and appreciates CHWs. CHWs engage with community structures.</td>
<td>Provide incentives/recognition</td>
</tr>
<tr>
<td></td>
<td>Ongoing data-based dialogue</td>
</tr>
<tr>
<td></td>
<td>Use of community influencers</td>
</tr>
<tr>
<td></td>
<td>Community structure engagement</td>
</tr>
<tr>
<td>8. Opportunity for Advancement</td>
<td>Potential for advancement</td>
</tr>
<tr>
<td>How CHWs are provided career pathways.</td>
<td>• Project, government, community</td>
</tr>
<tr>
<td>Level 3 requires:</td>
<td>• Retention</td>
</tr>
<tr>
<td>Advancement is offered to CHWs, training opportunities are provided to learn new skills, and advancement rewards good performance.</td>
<td>• Percentage retained</td>
</tr>
<tr>
<td></td>
<td>• Length of service</td>
</tr>
<tr>
<td></td>
<td>• Reasons for leaving</td>
</tr>
<tr>
<td>9. Data</td>
<td>• Data collection tools</td>
</tr>
<tr>
<td>How community-level data flow to the health system and back to the community and how they are used for quality improvement.</td>
<td>• Feedback provided to community and local government</td>
</tr>
<tr>
<td>Level 3 requires:</td>
<td>• Data used for problem solving</td>
</tr>
<tr>
<td>CHWs document visits and provide data that are reported to public sector monitoring systems. Supervisors monitor data quality, and CHWs and communities use data in problem solving.</td>
<td></td>
</tr>
<tr>
<td>10. Linkages to the National Health System</td>
<td>• CM referrals</td>
</tr>
<tr>
<td>The extent to which the Ministry of Health has policies in place that integrate and include CHWs in health system planning and budgeting and provides logistical support to sustain district, regional, and/or national CHW programs.</td>
<td>• Formal health system recognition and support</td>
</tr>
<tr>
<td>Level 3 requires:</td>
<td>• Country ownership</td>
</tr>
<tr>
<td>Linkages between CHWs and the formal health system (Ministry of Health), including referral, recognition and appropriate CHW provisions.</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: AIM, Assessment and Improvement Matrix; CHW, community health worker; CM, community mobilizer.
of children, and in conservative Muslim communities, only women could enter another woman’s household. However, the project has realized that because men are often the decision makers about health care, the project should recruit male CMs or more married couples who can work together. Angola also found a need for male CMs for the same reason. In Kenya/Somalia, CM sex varies by setting. In urban areas 60% CMs are women, while in the sparsely populated northern arid counties and borders, only 11% are. The sex of the CM is context based and differs according to nomadic lifestyle, harsh terrain,

<table>
<thead>
<tr>
<th>Country</th>
<th>Sex (%) Women</th>
<th>Rationale</th>
<th>Literacy/Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angola</td>
<td>90%</td>
<td>Women preferred</td>
<td>Low literacy</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>89%</td>
<td>Community preference</td>
<td>55% with basic reading and writing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Religious beliefs</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Insecurity</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Women’s Development Army Leaders must be women by government policy</td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>97%</td>
<td>Women preferred</td>
<td>Basic high school education</td>
</tr>
<tr>
<td>Kenya/Somalia</td>
<td>29%</td>
<td>Community preference</td>
<td>Basic reading and writing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Religious beliefs</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Insecurity</td>
<td></td>
</tr>
<tr>
<td>Nigeria</td>
<td>99%</td>
<td>Community preference</td>
<td>Some literacy; value of literacy diminished over time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Religious beliefs</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Insecurity</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Influence in the community</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CGPP, CORE Group Polio Project; CM, community mobilizer; NGO, nongovernmental organization.

*Data from the Secretariat Directors as of August 2019.*
Community Mobilizers’ Role in Polio Eradication in Remote Settings

sparsely populated communities, and community preference.

Commonalities on basic selection criteria of CMs existed across all projects, including the following:

- Known and respected members of the community
- Willing and committed to the welfare of the community
- Free from bad or corrupt behaviors
- Willing and available to learn and work for little or no money

Role. Key CM roles in CGPP included community mobilization to increase polio and routine vaccination rates, community-based surveillance of AFP, data collection, and promotion of maternal and child health. The exceptions were India, which did not include AFP surveillance because the country already had a robust system, and Kenya/Somalia, where CGPP worked directly at health facilities because of staff shortages. All CGPP programs used a wide variety of social mobilization methods to increase polio and routine vaccination rates: household visitation, group counseling, and community activities to dispel rumors and build trust in the health care system. A 2019 evaluation found that CMs engaged community leaders, created relationships with influencers, worked with household caregivers, and changed community attitudes that yielded normative and community change, not just individual behavior change. By identifying and reporting suspected polio cases for later investigation, CMs increased AFP surveillance sensitivity.

Program strategies had to respond in a timely way to specific local challenges and culture. In India, tactics changed as local OPV attitudes shifted from early acceptance, to suspicion and resistance, followed by passive acceptance and growing apathy. The Box provides an example of the evolution of the CM’s role and tasks in India. Edutainment (street theatre and puppet shows) was often used by CMs, but locations shifted over time. In Ethiopia, community trust resulted from content messages addressing traditional beliefs about the spiritual causes of paralysis. In Nigeria, 1,200 community influencers and edutainment such as community clowns were important drivers of behavior change. Kenya/Somalia had CMs meet with community members wherever they gathered, such as at bus stops and schools. Data collection included community mapping and tracking immunization status/defaulters and newborns.

Although Angola added polio activities to existing duties of child survival CHWs, other countries initially focused solely on polio eradication. However, over time, communities complained that other pressing health needs were not being met, so CGPP began to address other maternal and child health issues, as well as water and sanitation through their CMs.

Workload. Table 8 shows that workload varied by country and within country depending on terrain, culture, population density, and community traditions. In Angola, Ethiopia, Kenya/Somalia, and Nigeria, work averaged between 2 and 4 days per week and from 2 to 5 hours per day worked. However, CMs were expected to be available full-time during polio campaigns and other special events. Median average hours per month for CMs ranged from 16 in Ethiopia to 80 in India, and median average households reached monthly ranged from 75 in Angola and Ethiopia to 450 in India. As shown in Table 8, median average for hours per month was directly related to the number of households for which a CM was responsible. Variations in coverage included pastoralist areas in Ethiopia, shared workloads in Nigeria, and changes in program strategies over time. In India, initial work was only 1 week/month during campaigns, but after 2003, the workload increased with social mobilization.

Broad social engagement required working with multiple sectors of the community, and each country developed new strategies.

Training

Training varied by country and situation, but all CGPP programs conducted initial training immediately after CMs were recruited. Angola had the longest training period (2 weeks) to prepare CMs as both child survival and polio workers. Ethiopia, India, Kenya/Somalia, and Nigeria had initial trainings from 3 to 5 days, which was sometimes residential. Training initially focused on polio and
BOX. Evolution of Roles and Tasks of Core Group Polio Program/India Community Mobilizers

Core Group Polio Program (CGPP)/India trained and supported community mobilizers (CMs) to engage and convince communities, especially mothers/caregivers, about the benefits of vaccinating their children repeatedly for polio and to ensure that families were motivated to vaccinate their children for other life-threatening diseases.

Initially, the CM’s primary task was to:

- Mobilize community participation at government-run polio booths—1-day events to immunize all children under-5 at a fixed site on fixed days
- Conduct follow-up visits with families who were missed during a vaccination event

However, in certain remote and underserved areas, community resistance developed to a polio-focused strategy primarily due to frustration with and distrust of government, religious fatwas by Muslim leaders, and frustration at the lack of health services resulting in high numbers of sick children. To overcome the resistance, CMs had to broaden their role:

- Conduct monthly house-to-house visits often over years
- Facilitate community group meetings (such as mothers’ groups, religious groups)
- Use key community sites such as mosques, schools, and festivals for polio-related education, counseling, and problem solving
- Develop detailed maps of their communities and identify houses with unvaccinated children and later newborns
- Maintain immunization status records for all under-5 children in their areas
- Recruit community, cultural, and religious leaders to accompany them during household visits and to act as credible communication sources to dispel fears and rumors

As CMs learned more about community concerns, CGPP built CM capacity in counseling and in how to use community-relevant training materials and job aids.

The CM’s role grew to include:

- Promoting a larger package of services that responded to community needs and underlying causes of polio transmission, including routine immunization, water and sanitation, control of diarrheal disease, and breastfeeding.
- Referring and accompanying families to health facilities, building trust in the health system.
- Reaching migratory communities that had limited access to information or health services and were at risk of spreading the virus
  - Identifying and training key informers, such as barbers, employers, shopkeepers, and others, who knew the location and movement of migrant families in their area.
  - Developing maps of these populations and vaccine-eligible children, updating the maps regularly with socioeconomic information, and making regular visits to the mobile sites.
  - Forwarding information to government immunization teams to come to the sites to vaccinate the children.

immunization, community/social mobilization and interpersonal communication, but in each country the content expanded to additional maternal and child health topics. Although CGPP and NGO staff were the primary trainers, the curriculum was developed with government, UNICEF, and WHO staff as well as experts from health facilities, local governments, and/or international agencies.

Trainings evolved from lectures and presentation into very participative and creative sessions with role playing and household visits, flipbook messaging, dealing with body language, arguments, showing respect, and active listening. Continuous training occurred in all countries through supportive supervision, monthly CM meetings in India, Ethiopia, and Nigeria, and annual CM meetings in India and Kenya/Somalia. The first final evaluation recommended that CGPP update and strengthen its CM curriculum, and expand CM capacity in Ethiopia and India, by increasing the frequency of refresher training.

**Accreditation**
In each CGPP program, the supervisor periodically assessed health knowledge and competencies but no certification system was in place. However, each country program had an outside Knowledge, Practice, and Coverage evaluation of the CMs, verified by community recollection of receiving CM messages and support. For example, in India in 2011, 97% of mothers with a child 12–23 months knew their CM, and homes visited by CMs had higher levels of routine immunization than children whose homes were not visited.

**Equipment and Supplies**
All CGPP countries provided a continuous supply of job aids, which included flip books, registers, writing books, pens, posters, and sometimes bicycles.

**Supervision**
In each country, CGPP had a supervision system that reached upward to national and/or state government oversight. CGPP and project staff trained supervisors with expert and government input. The only variability between countries was the number of CMs supervised. This number ranged from 3–5 in Ethiopia to 12–15 in India. All project supervisors were paid or received a stipend. CGPP hired supervisors in all countries except for Ethiopia. In Ethiopia, the health extension workers (HEWs) are trained and supervised by the government; each HEW supervised CMs and 20–35 volunteer WDA leaders. India developed a 4-day master training course for supervisors. All projects used supervisory checklists and several types of data registers (immunization status records, pregnancy tracking, households visited).

The first final evaluation for 1999–2007 recommended more supportive supervision to address CM performance gaps. By 2017, supervision across projects was consistent, with supervisory checklists, registers, on-the-job visits, monthly and quarterly meetings, and performance evaluations in all countries.

Supervision tools, forms, templates, and training manuals can be found at the CORE Group website: https://coregroup.org/polio-eradication-toolkit/.

None of the CMs in any country received a regular salary. However, India, Nigeria, and Kenya/Somalia (except for urban areas) provided a monthly honorarium (average $30–$35), which enhanced motivation but sometimes created discontent when other projects provided a higher monthly amount. The Indian CGPP secretariat helped its mostly female cadre open bank accounts and deposited the funding electronically to ensure that the women had more control over their money. Angola, Ethiopia, Kenya/Somalia, and Nigeria provided a daily transport and food allowance (cash or food) during campaign days or

<table>
<thead>
<tr>
<th>TABLE 8. CORE Group Polio Project Community Mobilizer Workload, by Country Program</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days of Work and Household Coverage</td>
</tr>
<tr>
<td>Average days per week</td>
</tr>
<tr>
<td>Average hours per day/worked</td>
</tr>
<tr>
<td>Median average hours/month</td>
</tr>
<tr>
<td>Median average households/month</td>
</tr>
<tr>
<td>Range of households reached monthly</td>
</tr>
</tbody>
</table>
CM meetings. In Angola, this allowance was paid by the government.

A mix of nonfinancial incentives evolved over time in all countries. These were provided at 3 levels: country program, government, and community. Each country program was encouraged to introduce new incentives to boost morale and celebrate CM achievements. All provided CM training and skill development, which was highly motivational. All programs provided branded uniforms to identify and motivate the CM. Items included polio branded t-shirts, gowns, aprons, wrap-around skirts, umbrellas, rubber boots, coats, bags, streamers, caps, and scarves. In Nigeria, the CMs were given pink hijabs and they became known as the “pink ladies.” Certificates of recognition were given to CMs in Ethiopia, India, and Kenya/Somalia. Government recognition included officials providing CMs with formal recognition at public events in India and Nigeria. Ethiopia, India, and Nigeria gave performance-based awards to CMs, including mobile phones, radios, and shoes in Ethiopia; plaques in Nigeria; and trophies in India. Kenya/Somalia provided free access to health services for CMs and their dependents.

Community recognition was especially important to CMs and program staff in all the countries. It was considered a major motivator and essential to the work of CMs. In Kenya/Somalia, CMs were seen as community resource people and were invited to participate on committees. In Nigeria, some parents named their babies after the CMs. In India, CMs were recognized at community jamborees, where CMs received gratitude from the community for their referrals and other services to families. Ethiopia CMs were recognized by providing a certificate for their participation and contribution as volunteers.

Community Involvement

As previously described, the community was involved in every aspect of supporting the CM including recruitment and selection; feedback; problem solving, especially related to resistant households; incentives and recognition and ongoing dialogue about polio data.

Four of the programs (all but Angola) had a deliberate strategy of identifying and training community influencers to support the CMs. The community influencers helped solve problems with resistant households and monitored the effectiveness of the CM. Mothers’ groups met to discuss their polio indicators compared with other areas. Children were mobilized to be campaign advocates.

Ethiopia worked with religious and clan leaders, traditional healers, and ward (kebele) leaders to counter rumors. In Nigeria, as the project evolved, men were trained as peer informants for polio education and advocacy. In all countries, communities extended appreciation and respect for the CMs.

The content discussed in trainings and at supervision events emphasized concerns of the local community.

Opportunity for Advancement

Little potential existed for CM advancement within projects because the programs moved frequently to contain the virus. However, exceptional CMs had an opportunity to advance to the next level in India and Nigeria. CMs could become part of the CHW government cadre in India, Ethiopia, and Kenya. In many countries, because CMs were respected, they were invited to other community positions.

Ethiopia, India, and Nigeria all had very high retention rates (86%–95%) in difficult areas (there was no increase in compensation for this work). In Nigeria, experienced CMs delivered more CM messages during household visits than less experienced CMs. In Kenya/Somalia, retention was only 40% due to the nature of pastoralist communities crossing borders to follow herds and men finding paying jobs.

Data

CGPP commissioned 3 external final evaluations (1999–2007, 2007–2012, 2012–2017), each of which had recommendations for improving program performance. Initially, CGPP reported on achievements using primarily quantitative counts of activities with supporting anecdotes. In 2008, financial support was provided for baseline household surveys in local service areas and each country added a monitoring and evaluation officer.

In all countries, CMs maintained community maps, registers of pregnant women and newborns, defaulters, child immunization status, and households visited. These maps and registers were shared with health facility staff and next levels of government during supervision visits. Feedback was provided at community meetings and during local government meetings in countries such as Ethiopia and Nigeria. In India, data were posted on a community board. Copies of
registers for each country can be found at https://coregroup.org/polio-eradication-toolkit/. In addition, CGPP carried out household surveys in all countries to verify information. In India, CGPP used a census-based management information system for collecting prospective and retrospective information for planning, monitoring, and evaluation of its social and behavior change communication activities. India also used Lot Quality Assurance Sampling surveys and Barrier Analysis to identify impediments to adopting healthy behaviors.  

**Linkages to the Health System**

**CM Referrals.** All countries had a referral system for polio, and most programs evolved over time to provide referrals for routine immunization, antenatal care, newborns, and childhood and adult illnesses. In all countries, referral was viewed as one of the most important CM activities and one that led to community recognition. CMs were guides for vaccination teams to household defaulters. In many countries, CMs accompanied clients to the nearest health facility. In Kenya/Somalia, CMs guided pastoralists to the nearest health facility once they crossed the border, and they reported animal health issues to health facilities and veterinarians. In Angola, CMs gave caregivers referral slips for the health facility that could then be tracked, allowing the project to assess client follow-through.

**Formal Health System Recognition and Support.** CGPP programs worked to connect with health facilities, health workers, and government agencies in all countries. When CGPP started in Angola, the project actually supported the health system because of conflict and limited government functionality. When the program started, the government’s community health approach did not include accredited social activists (ASHAs). Auxiliary nurse midwives (ANMs) were the vaccinators, and the CMs guided them to resistant households during supplementary immunization activities. Over time, the program established coordinating meetings under the ANM with anganwadi workers (focussed on food supplements and nutrition), ASHAs, and CMs.

When CGPP began in Ethiopia, NGOs were working with CHWs for child survival projects, although all were using different approaches. In 2003, the government began its HEW program training full-time salaried women for 1 year. In 2005, 35,000 HEWs were deployed to work, with 2 HEWs stationed at each health post serving about 5,000 people. CMs became HEW guides in the community, and they planned together at the health facilities. In 2012, the government introduced the volunteer WDA to promote health, education, and agriculture. Some CMs became WDA leaders. In areas without WDA Leaders, CGPP continued to work with its CMs. CGPP found that WDA volunteers, given their other tasks, were not as effective at polio work as polio-focused CMs.

CGPP in Nigeria coordinates with the government structure at the federal, state, and LGA local (ward) levels. The program responds to calls from the government for additional support. For example, CGPP participated in a massive 8-day inactivated polio vaccine campaign in 2014 in the conflict-affected zones of Borno and Yobe states, the largest polio campaign in a conflict area in Africa, reaching nearly 800,000 children.  

In Kenya/Somalia, CGPP attached itself to health border facilities and established cross-border health committees primarily to address polio but also to address outbreaks of cholera and other diseases. CGPP provides transportation to health facility staff for immunization outreach targeting high-risk mobile populations along the border. In Kenya, community health extension agents, community health assistants, and health facility staff help monitor the CMs.

When new government cadres were introduced, some tension with CMs arose, but it was reduced with collaboration and communication over time.

When CGPP started, donors and governments thought polio would be quickly eradicated, so health ministries did not anticipate long-term investment in polio CMs. In countries that ended polio work, such as Angola, CMs were not absorbed into a government system. However, after 20 years of CGPP polio work in these countries, country governments are developing plans to integrate some of the CMs into their strategic health plans.

**DISCUSSION**

CGPP has worked with populations that were often resistant to immunization and required multiple doses of OPV. These populations were in difficult-to-access places due to conflict, rough terrain, and lack of roads. In addition, health services, staff, and infrastructure were limited, and people spoke multiple languages within countries and followed different tribal customs. Eventually in some areas, CGPP found it necessary to establish
Community involvement was the bedrock of the program and the original impetus for having CGPP work in remote settings.

cross-border programs with multiple immunization sites at various border-crossing points.

Although many countries hired part-time community workers to assist in polio campaigns, all CGPP countries built the capacity of community residents to volunteer part-time all year long and to make house-to-house visits. The CGPP CM strategy, based on CHW work in child survival programs, evolved over time. Initially, Angola added polio interventions onto existing CHW tasks. Ethiopia focused primarily on surveillance for AFP, and India focused on social mobilization. Learning took place from technical assistance visits from CGPP headquarters and secretariat staff, and later from a set of midterm and final evaluations, conferences and workshops, and a growing literature on CM best practices. Learning was eventually codified into templates, curricula, and articles that benefitted more recent projects in Nigeria and Kenya/Somalia. Each country adapted its CM practices to respond to local needs. This adaptation reflects similar findings that the CM evidence base needs to be contextualized for different places and situations.20

All AIM components, with the exception of accreditation, were used. Three were partially met and 6 achieved basic functionality (Table 9). Of the set of 10 components, data, both quantitative and qualitative were critical to program performance. In 2008, CGPP developed community-based health information systems and added a monitoring and evaluation officer to each country. Indicators moved from counting activities, houses, and people to providing useful household information aggregated at community and district levels around key global polio and immunization indicators that improved decision making. A focus on supervision systems, robust data collection systems combined with periodic surveys, and performance assessments created functioning CM systems over time. Strategic use of data at every level, from planning and message development to results monitoring, was previously highlighted as one of the 4 lessons learned from the CGPP/UNICEF India experience.21

CGPP Nigeria rapidly improved its OPV 0 dose by asking its CMs to attend “naming ceremonies” since so many women gave birth at home. CGPP Ethiopia increased its OPV 3 coverage to double that of non-CGPP areas within the same state by focusing on messaging to lower the drop-out rate.2 CGPP India increased its full routine coverage in CGPP catchment areas of Uttar Pradesh from 48% (2008) to 78% (2017), well above Uttar Pradesh’s state coverage (51% in 2016) by using local data from CGPP’s census-based management information system.3 India conducted barrier analysis to examine factors responsible for timely OPV 3, finding that respondents who perceived other benefits of child immunization were 3 times more likely to timely vaccinate their children than those who did not, allowing them to adapt their messaging.22 Each country commissioned Knowledge, Attitude, and Practice surveys of the CMs and community residents that provided data for new messaging, refresher training, and increased supervision. High-quality information and data analysis increased credibility of CGPP with the formal health system over time. The creation and implementation of high-quality data collection, analysis, and the utilization of the information for program improvement was a major contribution of CGPP. These steps are essential for ongoing routine immunization efforts in all countries.

Community involvement was the bedrock of the program and the original impetus for having CGPP work in remote settings. Communities were integrated throughout several AIM components in which functionality was met. Other research has found that CHWs are embedded in the community when community members trust and respect them and feel a sense of ownership over the program.1 CMs were from the same community and ethnicity, identified by the community, and community members contributed ideas and supported CMs in their tasks. CMs did not work in isolation; in all cases examined, they required the support of community leaders and influencers. This need for support was especially true in areas where trust in government health programs was low or families were vaccine-hesitant. The CGPP strategy depended on CMs providing behavioral change visits to households most at risk of missed vaccination, using detailed community maps. Evidence supports the conclusion that these house-to-house visits played a role in increasing OPV 3 completion in Ethiopia and India.6 In 2013, the Global Polio Eradication Initiative made a new strategic plan, which recognized that the conventional eradication strategy needed to be supplemented by efforts to increase community participation according to local needs for a multi-pronged, area-specific strategy that would vary in different settings.23

Communities participated in the CM role definition and recruitment, training, appropriateness of equipment and supplies, and incentives (community recognition was a major support). The sex of the CM was dependent on the community’s
<table>
<thead>
<tr>
<th>CHW AIM 2018: Revised Programmatic Components</th>
<th>CGPP Achievement of Level 3 Functionality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Role and Recruitment How the community, CHW, and health system design and achieve clarity on the CHW role and from where the CHW is identified and selected.</td>
<td>Level 3 achieved Clarity and clear criteria identified for recruitment and role. Some criteria changed over time.</td>
</tr>
<tr>
<td>2. Training How preservice training is provided to CHWs to prepare for their roles and ensure they have the necessary skills to provide safe and quality care. How ongoing training is provided to reinforce initial training, teach CHWs new skills, and help ensure quality.</td>
<td>Level 3 achieved Initial training in 4 of 5 programs 3–5 days, maximum was 2 weeks in Angola. Trainers included CGPP and NGO staff with health facility and government officials and other resource people varying. Training content in addition to polio, provided broad maternal and child health and social and behavior change skills in most programs. On-the-job mentoring was the major method of continuing education through CM mentoring, monthly meetings, and annual meetings.</td>
</tr>
<tr>
<td>3. Accreditation How health knowledge and competencies are assessed and certified prior to practicing and recertified at regular intervals while practicing.</td>
<td>Level 3 not achieved because there was no formal certification system. CM health knowledge and competencies assessed initially and periodically. External program evaluations documented Knowledge, Practice, and Coverage of CMs and verified with community.</td>
</tr>
<tr>
<td>4. Equipment and Supplies How the requisite equipment and supplies are made available when needed to deliver expected services.</td>
<td>Level 3 achieved Continuous supply of job aids (flip books, registers, writing books, pens, posters, sometimes bicycles).</td>
</tr>
<tr>
<td>5. Supervision How supportive supervision is carried out such that regular skill development, problem solving, performance review, and data auditing are provided.</td>
<td>Level 3 achieved All country programs addressed supervision at all levels and types.</td>
</tr>
<tr>
<td>6. Incentives How a balanced incentive package reflecting job expectations, including financial compensation in the form of a salary and nonfinancial incentives, is provided.</td>
<td>Level 3 partially achieved CMs were part-time workers and did not receive a salary. In 3 of 5 programs, CMs received a monthly honorarium (underpaid compared to UNICEF). All provided transport/food allowances for campaigns and program meetings. 3 of 5 provided certificates and performance awards. 1 provided free access to health services. All had community recognition.</td>
</tr>
<tr>
<td>7. Community Involvement How a community supports the creation and maintenance of the CHW program.</td>
<td>Level 3 achieved This was one of the strongest components of the CGPP. All programs demonstrated strong and continuous community involvement.</td>
</tr>
<tr>
<td>8. Opportunity for Advancement How CHWs are provided career pathways.</td>
<td>Level 3 partially achieved Because the program was vertical and had changing geographic areas, opportunities to advance within the program were limited. 3 of 5 programs reported opportunities in government and community. Retention was high (86%–95%) in 3 of the programs, 40% in another, no data in the fifth.</td>
</tr>
<tr>
<td>9. Data How community-level data flow to the health system and back to the community, and how they are used for quality improvement.</td>
<td>Level 3 achieved Data collection tools included community maps, registers of pregnant women and newborns, defaulters, child immunization status, and household visits. Feedback was provided to community and local government and health system. Data were used for problem solving to improve program performance.</td>
</tr>
<tr>
<td>10. Linkages to the National Health System The extent to which the Ministry of Health has policies in place that integrate and include CHWs in health system planning and budgeting and provides logistical support to sustain district, regional, and/or national CHW programs.</td>
<td>Level 3 partially achieved Because program was vertical and had limited time expectations, it was never fully integrated with the national health system even though CM referrals were made and CMs worked closely with government cadres in all countries.</td>
</tr>
</tbody>
</table>

Abbreviations: AIM, Assessment and Improvement Matrix; CGPP, CORE Group Polio Project; CHW, community health worker; CM, community mobilizer; NGO, nongovernmental organization.
religious and cultural preferences. Women were often selected because of their role in family caregiving or access to households, especially in Muslim communities. Men were selected because of geographic and population challenges such as harsh terrain, pastoral movement, limited phone networks, or conflict and insecurity. All programs recognized the need to include both men and women for social mobilization and decision making to improve vaccination and other child health indicators, and they developed community influencer strategies to support CM work. The roles of CMs expanded over time to include perceived needs of the community (e.g., water, sanitation, roads, antenatal care, newborn care, injuries). Validation of community-defined issues and responsiveness to them improved vaccination rates and demonstrated the need for integrated services in vertical programs, similar to previous findings. In countries where community-based surveillance of AFP was prominent, non-polio AFP rates expanded in CGPP communities, often exceeding the national rate (see Table 1). In addition, the community-based AFP system allowed for rapid identification and response to other disease outbreaks such as chikungunya and measles.

Linkages to the health system component were only partially functional. Because CGPP was a vertical program with limited time expectations, it was never fully integrated into the national health system. However, the support of government and the ministry of health at the national and state level and the strong linkage of the CM program to local government leaders and committees were essential to polio eradication in these hard-to-reach communities. Referrals to health facilities for immunization and other illnesses were viewed by CMs as one of their most important duties, which was an incentive. The partnership worked both ways. Influential government leaders reduced some of the rumors and hesitancy and provided the CMs with recognition. CGPP trained government health workers to supervise and utilize the skills and local knowledge of the CMs. Ludwick et al. noted that factors pertaining to supportive supervision and relationships with other health care workers related to variances in performance outcomes within a program.

CGPP’s country programs demonstrate how a CHW model can be utilized in a vertical program and adapted to meet specific country and community needs. Other research identified 4 essential elements for an enabling CHW work environment that were also found in this project: workload, supportive supervision, supplies and equipment, and respect from the community and the health system. The CHW AIM Tool proved useful in systematically assessing CHW functionality of a vertical program. Others have found it useful as a participative exercise with village health teams in Uganda for integrated programs.

CONCLUSION

This cross-country analysis of the CORE Group Polio Program’s Community Mobilizers demonstrates the importance of the full range of AIM components, even in a vertical program. It also suggests that vertical programs need to expand to address community needs if they are to be effective in meeting their original goal. Data, including local registration of vital events and child registries, played a critical role in program improvement and constitute an essential component. Community engagement is also critical to address misinformation, vaccine hesitancy, and mistrust of government—such engagement needs to be tailored to each culture and community. Community-based surveillance using local volunteers, especially in hard-to-reach populations, enhanced national and state efforts. Partnerships and communication with government health systems are important for program credibility, success, and sustainability.

These lessons are important at this point in time because of the variety of vertical programs and disease challenges from measles, Ebola, and the current COVID-19 pandemic, as well as non-communicable diseases. There is value in using a similar approach to that used by CGPP and its CMs for responding to COVID-19, as well as other global public health priorities. Responses to COVID-19 should engage the community through its community mobilizers for nuanced and repeated messaging and discussion to improve the knowledge and attitudes of different community groups about the virus and to keep their trust. The CM’s role can add value to government efforts on disease prevention, testing, contract tracing, home visiting, and community support. Once a vaccine is developed, CMs could mobilize communities for high vaccination coverage.

Acknowledgments: We would like to acknowledge the contributions of the CORE Group Secretariat staff from Angola, Ethiopia, India, Kenya/Somalia, and Nigeria and the headquarters CORE Group Polio Project staff.

Competing interests: The first 2 authors received payment from the CORE Group Polio Project Secretariat based at World Vision USA. The other authors are employees of CORE Group Polio Project in their respective countries.
REFERENCES


Integrating Calcium Into Antenatal Iron-Folic Acid Supplementation in Ethiopia: Women’s Experiences, Perceptions of Acceptability, and Strategies to Support Calcium Supplement Adherence

Gina C. Klemm, Zewdie Birhanu, Stephanie E. Ortolano, Yohannes Kebede, Stephanie L. Martin, Girma Mamo, Katherine L. Dickin

Key Findings

- In a small acceptability trial of global antenatal calcium supplementation guidelines, women preferred 2-event regimens to 3- or 4-event regimens, but acceptability was not associated with higher adherence.
- Rural Ethiopian women were motivated and adhered to antenatal micronutrient supplementation despite regimen complexity, forgetfulness, side effects, and discouragement from others.
- Simple home-based strategies and family support contributed to high adherence.

Key Implications

- Formative research can guide antenatal micronutrient supplementation programs by identifying adherence strategies that reflect women’s views and address misperceptions.
- With appropriate program design, supportive counseling, and sufficient access to supplements, women can be highly motivated to adhere to antenatal micronutrient supplementation.

ABSTRACT

Recommendations for antenatal calcium supplementation to prevent preeclampsia could substantially reduce maternal mortality, but adherence to multiple daily doses may constrain effectiveness. World Health Organization guidelines recommend 3 daily calcium supplements (1.5–2 g/d), taken separately from 1 iron-folic acid (IFA) supplement; however, limited data suggest lower calcium doses may also be effective. We conducted mixed-methods household trials to identify strategies for supporting adherence and integrating calcium into antenatal IFA supplementation programming in Ethiopia. Participants were randomly assigned to 3 regimens varying in dose and timing and were later given a choice of regimens. Semistructured interviews conducted over 6 weeks explored acceptability, barriers, and facilitators and offered opportunities to choose calcium pill type. Interviews were transcribed, translated, and analyzed thematically. Calcium adherence was measured using medication event monitoring. All participants (N=48) agreed to try supplementation. Adherence barriers included forgetting to take pills when busy or travelling and perceived side effects. Midday doses were the most challenging because of farming, market, and social events; women avoided taking supplements in public due to fear of being perceived as HIV positive. Social support from families, visual reminders, and anticipated benefits motivated adherence. More participants (75%) selected chewable versus conventional supplements due to organoleptic properties, but this preference declined over time. Adherence rates did not substantially differ across regimens with 2 (81.1%), 3 (83.4%), or 4 (77.1%) pill-taking events. Women indicated that the 2-event regimen was more acceptable than 3- and 4-event regimens, but this acceptability was not associated with higher adherence. Consequently, mean daily calcium consumption (811.3 mg) was lower than for 3-event (1,251.1 mg) and 4-event (1,156.4 mg) regimens. Integrating calcium into antenatal IFA supplementation is acceptable to Ethiopian women, with a 3-event regimen yielding the highest consumption rates. Despite women experiencing challenges with midday dosing and stigma, using simple home-based strategies and being counseled on the purpose of supplementation were more effective than reducing dosage for mitigating barriers and improving adherence.

a Program in International Nutrition, Division of Nutritional Sciences, Cornell University, Ithaca, NY, USA.
b Faculty of Public Health, Department of Health, Behavior and Society, Jimma University, Jimma, Ethiopia.
Correspondence to Katherine Dickin (kld12@cornell.edu).
INTRODUCTION

Antenatal micronutrient supplementation is a cost-effective, scalable approach that can contribute to addressing the persistent challenge of maternal undernutrition1–3; effective and broad implementation of interventions with proven efficacy is urgently needed.4,5

Global guidelines on antenatal calcium supplementation to prevent preeclampsia, a leading cause of maternal mortality, have yet to be widely adopted into national policies.6,7 A systematic review of efficacy trials found that calcium supplementation in pregnancy reduced the risk of preeclampsia by half.8 Such trials critically demonstrate the potential impact of micronutrient supplementation, but they provide limited guidance for programming in different settings.9

To prevent preeclampsia, the World Health Organization (WHO) suggests 1.5–2 g of calcium per day, taken in 3 divided doses with food and separately from 1 daily iron-folic acid (IFA) supplement.7 IFA supplementation to prevent anemia is a longstanding example of a global guideline that has been translated into nutrition policies and programs in many country contexts.10 Intake of antenatal IFA is often inadequate in Africa and Asia due to problems of limited coverage and supply, as well as barriers to adherence.11 Rates of IFA consumption vary substantially across countries in sub-Saharan Africa, with adherence related to income, education, and initiation and frequency of antenatal care (ANC).12 Pooled national antenatal IFA consumption rates in Ethiopia are less than 50%, with much lower rates in rural areas; persistent adherence barriers include lack of first trimester ANC use, inadequate counseling, limited knowledge of supplements and anemia, fear of side effects, forgetfulness, and low adherence support.13–15 Although some reports define “adherence” as consumption of IFA supplements for ≥90 days during pregnancy, it is important to recognize that consumption may be limited by systemic supply constraints (including costs of national procurement and distribution) on women’s access to supplements.11,16,17

A regimen including calcium as well as IFA is likely to exacerbate implementation and adherence challenges well known within IFA programs.11,16,17 Research is needed on how best to integrate calcium supplementation into IFA and ANC programming, given that strong supplementation programs are still lacking in many contexts despite decades of IFA advocacy and experience.18–20

In related formative studies in Kenya and Ethiopia, researchers explored how lessons from IFA supplementation can inform calcium initiatives and reduce well-known barriers.21,22 Additional research has examined the influence of regimen complexity or product type on adherence to calcium23 and IFA,24 as well as the feasibility of calcium supplementation delivery through the health system.25–27 Evidence on acceptability of products, regimens, and counseling messages can inform programs integrating calcium and IFA supplementation and help to maximize utilization.

We aimed to assess factors that influence acceptability and adherence to calcium supplementation among pregnant women in Ethiopia, including supplement type, regimen complexity, administration with IFA, and emergent barriers and facilitators. To assess the acceptability of integrating calcium and IFA supplementation, we used trials of improved practices (TIPs), a mixed-methods approach that tests proposed health interventions and behavioral recommendations to ensure that intervention design is locally acceptable.28,29 During multiple home visits, interviewers and participants discuss the latter’s willingness to try new behaviors, what is difficult or easy about executing selected behaviors, and modifications to improve acceptability.

In addition to qualitatively exploring the motivations, preferences, barriers, and facilitators that influenced acceptability of calcium supplementation, we compared women’s adherence across 3 possible dosing regimens. Previous research found that the number of prescribed daily doses and the regimen complexity were inversely related to adherence of oral medication regimens.30–33 We hypothesized that the regimen suggested by WHO (3 divided doses daily and separate administration of IFA) could result in low adherence rates. In light of evidence that any negative effects of calcium on iron absorption are short-term and calcium supplementation does not adversely affect iron status over time,34 we tested a regimen that permitted calcium and IFA to be taken together. Additionally, because evidence suggests that lower calcium doses may be effective,35,36 we tested an alternative lower-dose regimen (1 g/d). We hypothesized that women might find these alternative regimens with fewer daily administrations more acceptable, resulting in higher adherence rates and comparable consumption to the WHO recommended regimen (1.5–2 g/d).

We analyzed women’s perceptions and experiences, as well as adherence rates, to inform policy and program recommendations for translating...
global guidelines on antenatal calcium supplementation into ongoing ANC programs within Ethiopia and internationally.

**METHODS**

**Design**

This mixed-methods study used the TIPs approach to conduct a household-level exploration of women’s perceptions and experiences with a new recommended behavior—consumption of antenatal calcium supplementation in addition to IFA. During home visits, women were provided with supplements, counseled on their use, and then asked to participate in a series of in-depth interviews on their experiences and views of acceptability.

In addition, consumption of calcium supplements was assessed quantitatively to triangulate with women’s perceptions and reports and to compare rates of adherence and consumption across 3 regimens. We randomly assigned women to 3 regimens (Table 1) that integrated calcium into the IFA supplementation regimen recommended in Ethiopia. These represented the calcium regimen proposed in the WHO guidelines (4-event; 1.5 g of calcium), as well as a simplified regimen allowing co-administration of calcium and IFA (3-event; 1.5 g of calcium) and a lower-dose regimen allowing co-administration and reducing calcium dosage (2-event; 1 g of calcium).

**Setting and Participants**

This study was conducted in 2 rural districts in Oromia region, Ethiopia, from December 2014 to March 2015. In this area, most women (81%) deliver at home, half do not receive ANC, less than one-third (30%) take any iron tablets, and only 35% are informed of pregnancy danger signs. Limited awareness and minimal diagnostic tools and resources for preeclampsia/eclampsia suggest that many women do not receive adequate care. To assess the acceptability of calcium in the context of functioning IFA supplementation, we selected rural communities where a recent program, implemented by Nutrition International, had strengthened ANC and IFA supplementation.

In each of 2 districts, we purposively selected 5 health posts for diversity in proximity to a referral health center and location in rural and semi-urban communities. We included healthy pregnant women aged 16 or older who were between 12 and 28 weeks’ gestation at enrollment and who would be available for interviews scheduled over the 6-week study period. Exclusion criteria were high-risk pregnancy, high habitual calcium intake, intellectual disability, and severe illness.

Health extension workers at each health post were informed of the study purpose and asked to identify all pregnant women in the catchment area who had and had not attended ANC. Based on health registries and their knowledge of pregnant women in the community, the workers provided information on demographic characteristics to facilitate purposeful selection of women representing variation in ANC use, gravidity, education, and daily hours spent outside the home. Purposeful sampling is widely used in qualitative research to select information-rich cases and identify samples that include key characteristics expected to influence experiences and perspectives of respondents.

Selected women were screened using a questionnaire that included a brief 1-week food frequency measure focused on high-calcium foods. A nurse accompanying the study team followed standard antenatal history protocols to interview

**TABLE 1. Integrated Antenatal Calcium and Iron-Folic Acid Dosing Regimens Assessed in Household Trials in 2 Rural Districts, Ethiopia (N=48)**

<table>
<thead>
<tr>
<th>Regimen</th>
<th>Calcium Supplement&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Dosing Schedule</th>
<th>Iron-Folic Acid Supplement&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Dosing Schedule</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-event</td>
<td>3 (1,500 mg) Morning, midday, evening (with meals)</td>
<td>1 Before bed (separate from calcium)</td>
<td>1 Evening (with calcium)</td>
<td></td>
</tr>
<tr>
<td>3-event</td>
<td>3 (1,500 mg) Morning, midday, evening (with meals)</td>
<td>1 Evening (with calcium)</td>
<td>1 Evening (with calcium)</td>
<td></td>
</tr>
<tr>
<td>2-event</td>
<td>2 (1,000 mg) Morning, evening (with meals)</td>
<td>1 Evening (with calcium)</td>
<td>1 Evening (with calcium)</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> 500 mg elemental calcium.

<sup>b</sup> 65 mg elemental iron and 0.4 mg folic acid.

<sup>c</sup> Event is the number of discrete times per day that the regimen prescribes taking any supplement.
women and estimate gestational age based on last menstrual period. If eligible, women were invited to participate in the study, with invitations continuing until 5 women were enrolled in each of 10 communities. Five women were excluded after screening: 3 had an intellectual disability or serious illness, 1 was in the first trimester of pregnancy, and 1 planned to travel and would not be available to complete the study. No women were excluded for high calcium consumption or high-risk pregnancy.

For the quantitative comparison of adherence to calcium supplementation across the 3 regimens (Table 1), we estimated that a sample size of 15 women per group (N=45) would allow detection as significant a difference in intake between groups equal to 1 standard deviation (i.e., effect size = 1) with a 2-sided test, 80% power, and alpha of .05. We recruited 50 women to allow for 10% loss to follow-up.

**Intervention Materials**

Based on formative research, we knew that women often felt they received inadequate information about why and how to take antenatal supplements. We adapted counseling cards from IFA materials designed by Nutritional International-Ethiopia and developed a reminder calendar for women to use at home, with motivating messages and illustrations indicating how many supplements to take and when to take them.

To allow for personal preferences, we offered a choice of 2 calcium carbonate products (conventional and chewable). The conventional (hard) calcium supplement was Ostocal D (Eskayef Bangladesh Ltd.), which contained 500 mg of elemental calcium and 200 IU of cholecalciferol. The chewable calcium supplement was Ideas (Innothera Chouzy Ltd.), which contained 500 mg of elemental calcium and 400 IU of cholecalciferol. Women who did not have
IFA from ANC visits to health posts were provided supplements (Medicamen Biotech Ltd.) containing 60 mg of elemental iron and 400 μg of folic acid.

**Data Collection Procedures**

Five trained interviewers recruited 50 women, who were approached at home and screened for eligibility. Participants were randomly assigned to 1 of 3 supplementation regimens (Table 1). Allocation was distributed evenly across community and interviewer.

Participants were offered daily calcium and IFA supplementation for 6 weeks and interviewed on their experiences every 2 weeks (range: 13–15 days), totaling 4 interviews (1–1.5 hours each) per participant. Interviews were conducted using semistructured interview guides that had been translated into Afan Oromo, back-translated, and pretested in the local context.

Table 2 summarizes the sequence of activities. During visit 1, interviewers assessed a woman’s ANC and IFA supplementation experience and counseled her on the benefits of supplements, regimens, side effects, and adherence strategies using counseling cards and reminder calendars. Conventional and chewable supplements were presented so that women could examine them and choose the preferred supplement type, discussing the reasons for their choice. Interviewers asked about women’s willingness to try supplementation for 2 weeks until the next interview and addressed any concerns. Counseling was the same for all regimens, except for the regimen-specific details on how often to take supplements. Women were also asked about their views on asking a family member or friend to remind and encourage their supplement use; results on acceptability of adherence partners are reported elsewhere.43

Follow-up visits 2, 3, and 4 involved interviews about supplementation experiences, concerns, challenges, and facilitators. At the end of each interview, women were asked if they were willing to continue with supplementation; their interview responses were tabulated to assess acceptability.44 At each visit, participants also had the option to choose either conventional or chewable calcium supplements for the next period and were provided with supplements; this allowed us to assess preference for calcium products.

At visit 3, after 4 weeks on a randomly assigned regimen, women were also offered the option of choosing whichever regimen they preferred to follow, and then asked to discuss their reasons. This discussion provided deeper insight into different women’s responses to the 3 regimens, and the final interviews on visit 4 focused on the women’s experiences with the regimens they chose. On visit 4, women were also asked if they were willing to continue supplementation after the study ended, and if so, they were given supplements to last the remainder of their pregnancy.

### Table 2. Sequence of Activities for Trials of Improved Practices Testing Acceptability of Antenatal Calcium and Iron-Folic Acid Supplementation in 2 Rural Districts, Ethiopia

<table>
<thead>
<tr>
<th>Visit 1 (Enrollment)</th>
<th>Visit 2 (2-wk follow-up)</th>
<th>Visit 3 (4-wk follow-up)</th>
<th>Visit 4 (6-wk follow-up)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intervention</strong></td>
<td><strong>Supportive</strong></td>
<td><strong>Supportive</strong></td>
<td><strong>Supportive</strong></td>
</tr>
<tr>
<td><strong>Regimen</strong></td>
<td><strong>Random assignment</strong></td>
<td><strong>Maintained</strong></td>
<td><strong>Choice</strong></td>
</tr>
<tr>
<td><strong>Data collection</strong></td>
<td><strong>Willing to try?</strong></td>
<td><strong>Continued? Willing to continue?</strong></td>
<td><strong>Continued? Willing to continue?</strong></td>
</tr>
<tr>
<td><strong>Acceptability response</strong></td>
<td><strong>Trials?</strong></td>
<td><strong>Continued?</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Qualitative</strong></td>
<td>• Anticipated challenges/facilitators</td>
<td>• Experiences</td>
<td>• Experiences</td>
</tr>
<tr>
<td></td>
<td>• Motivations</td>
<td>• Challenges/facilitators</td>
<td>• Challenges/facilitators</td>
</tr>
<tr>
<td></td>
<td>• Concerns</td>
<td>• Motivations</td>
<td>• Motivations</td>
</tr>
<tr>
<td></td>
<td>• Others’ reactions</td>
<td>• Concerns</td>
<td>• Others’ reactions</td>
</tr>
<tr>
<td></td>
<td>• Strategies</td>
<td>• Strategies</td>
<td>• Strategies</td>
</tr>
<tr>
<td><strong>Quantitative</strong></td>
<td><strong>Demographics</strong></td>
<td><strong>Adherence (MEMS)</strong></td>
<td><strong>Adherence (MEMS)</strong></td>
</tr>
<tr>
<td><strong>Abbreviation:</strong> MEMS</td>
<td><strong>Experiences</strong></td>
<td><strong>Experiences</strong></td>
<td><strong>Experiences</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Challenges/facilitators</strong></td>
<td><strong>Challenges/facilitators</strong></td>
<td><strong>Challenges/facilitators</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Motivations</strong></td>
<td><strong>Motivations</strong></td>
<td><strong>Motivations</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Concerns</strong></td>
<td><strong>Concerns</strong></td>
<td><strong>Concerns</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Others’ reactions</strong></td>
<td><strong>Others’ reactions</strong></td>
<td><strong>Others’ reactions</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Strategies</strong></td>
<td><strong>Strategies</strong></td>
<td><strong>Strategies</strong></td>
</tr>
</tbody>
</table>

Participants received calcium and IFA supplements and counseling on benefits, side effects, and adherence strategies.
Adherence to calcium over the previous 2 weeks was evaluated at visits 2, 3, and 4 using medication event monitoring system (MEMS; AARDEX Ltd., Zug, Switzerland) technology with a microelectronic chip registering the date and time of each bottle opening. Bottle openings were the proxy for intake of 1 supplement. Participants were instructed to (1) only open the bottle to take 1 supplement; (2) keep the bottle away from children or others who may open the bottle; and (3) take the bottle when traveling. IFA (and calcium) adherence was additionally captured via self-report. At each visit, women received 40–60 calcium supplements (i.e., an amount slightly exceeding regimen requirements), and those who did not have IFA from ANC also received 2 10-capsule IFA blister packs.

**Data Analysis**

Interviews were audio-recorded, transcribed verbatim, and translated into English by each interviewer. ZB reviewed each interviewer’s first audio recording during each interview round against the transcript for accuracy of translation and quality of interview technique, and held daily debriefs to discuss challenges to the protocol. ZB and YK reviewed all transcripts for clarity and completeness.

GCK, SEO, and SLM independently read and coded 16 transcripts using a grounded theory approach to develop initial code categories based on participant experiences with supplementation. Investigators jointly reviewed codes and developed a codebook to guide systematic coding of transcripts in Atlas.ti version 7 (Scientific Software Development GmbH). GCK and SEO independently coded 30% of transcripts to standardize coding. All authors discussed and resolved differences in coding and code definitions. Results were organized into matrices summarizing participant experiences, preferences, and key quotations to allow comparison across cases and within cases. Matrices and relevant output were shared with the research team to
review consistency of coding and reach consensus on interpretation. Qualitative results are summarized in the narrative, describing approximate proportions of respondents who mentioned a point to give a sense of salience. Lists of reasons begin with the most commonly mentioned. Similar topics were discussed across all time points, so interview results from all visits are summarized by theme, noting changes in experiences over time. Illustrative quotes are included in the text and in a table on barriers and facilitators.

MEMS data were configured using Powerview (MWV) and cleaned using SAS/STAT, version 9.3 (SAS Institute Inc.). Percent adherence (or “adherence rate”) was defined as the number of MEMS events divided by number of prescribed doses during the assessment period multiplied by 100, and capped at 100% (capping occurred for 11% of MEMS measurements). Daily consumption is reported in milligrams and was defined as the number of MEMS events multiplied by 500 mg, divided by days in assessment period. The Assessment period was defined as the number of days between interviews. MEMS events recorded on interview days were not included in the analysis to exclude events resulting from demonstration and practice of bottle opening. One-way analysis of variance (ANOVA) was used to compare adherence rates and calcium consumed (mg) across 3 regimens. Significance was set at $P < .05$. The post-hoc Tukey HSD test was used to determine which group means were significantly different from each other. Analyses were performed using SPSS 24.0 (IBM Corp.). MEMS data were triangulated with self-reported adherence to categorize women by adherence level and inform interpretation of interview data.48,49

Ethical Approval
Cornell University’s Institutional Review Board (1205003071), the Ethiopian Public Health Institute, and the Oromia Regional Health Bureau approved this study. All study staff completed training on research ethics, protocol, and informed consent. All women approached agreed to participate and gave written informed consent.

RESULTS
Forty-nine of 50 recruited women (98%) completed the 6-week study; 1 participant withdrew after 2 weeks for unknown reasons. Another woman completed the study but had incomplete data so was dropped from analysis. Results are reported for the final sample of 48 women with complete data.

Participant Demographics
All but 3 participants were married and mean gestational age was 20.5 weeks (range 13–28 weeks). The majority were rural; most households owned livestock (73%) and grew crops (71%). Most women routinely spent time away from home (77%); a third spent over 5 hours away daily. Women’s educational levels were very low, and food insecurity was relatively common (Table 3). Demographic characteristics were similar across groups assigned the 3 regimens, although the 2-event group tended to be more food secure (Table 3).

Antenatal Care and IFA Usage Before Study
At visit 1, about 70% of respondents had attended ANC once or twice during the current pregnancy, on average starting at 3.5 (range: 1–6) months; a few had attended more often and almost 20% had not attended. Three-fourths of respondents received IFA before study enrollment, with most taking IFA for <1 month. Several women mentioned side effects and methods to minimize them; only 3 had discontinued use. Despite programming in the study sites to increase access to IFA information, women wanted more specific details:

1 want to know the detailed importance for mother and fetus. The health worker said, “Take 1 pill daily and come back on your appointment” but this is not enough. —28-year-old woman, 4-event, very high adherence

Willingness to Try Calcium and IFA
Across visits, all 48 women were willing to try or to continue calcium and IFA supplementation. At visit 1, when asked about potential difficulties with supplementation, most women said they did not expect barriers, anticipating the health benefits would motivate them to consume supplements as counseled. Six women, despite their willingness to try supplementation, initially did not adhere. Their adherence improved following clarification on visit 2 to ensure they understood the schedule and that they could take a missed supplement when they remembered instead of skipping a dose.

Initially, more than one-fourth of the women anticipated 1 or more barriers, including forgetting during work or travel, having side effects or feeling ill, taking too many supplements, or not having water or food available. A few women
preferred to keep calcium in a discrete place for fear others would assume it was antiretroviral therapy.

When asked what might facilitate adherence, nearly half of the women suggested keeping supplements in a visible place, about one-fourth suggested taking calcium with meals, and 9 women suggested reminders from family members.

Counseling women on the need to take calcium was challenging because women were not aware of preeclampsia or hypertensive disorders of pregnancy. Most women knew of anemia, half referring to it as “low blood” and IFA as the tablet that “fills blood.” A few women initially expressed uncertainty about taking calcium and IFA together for what seemed to them to be opposing health conditions, and further counseling was required for these women. Women who understood supplementation as a preventive measure accepted the need to take both calcium and IFA.

### Table 3. Characteristics of Pregnant Women Who Completed Household Trials on Calcium and Iron-Folic Acid Supplementation in 2 Rural Districts, Ethiopia (N=48)

<table>
<thead>
<tr>
<th></th>
<th>4-Event Regimen n=15</th>
<th>3-Event Regimen n=17</th>
<th>2-Event Regimen n=16</th>
<th>Total Sample N=48</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age, y, mean (SD)</td>
<td>28.5 (4.5)</td>
<td>26.0 (6.6)</td>
<td>26.4 (6.0)</td>
<td>26.9 (5.8)</td>
</tr>
<tr>
<td>Gravidity, No. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primigravida</td>
<td>4 (26.7)</td>
<td>5 (29.4)</td>
<td>4 (25.0)</td>
<td>13 (27.1)</td>
</tr>
<tr>
<td>Multigravida</td>
<td>11 (73.3)</td>
<td>12 (70.6)</td>
<td>12 (75.0)</td>
<td>35 (72.9)</td>
</tr>
<tr>
<td>Educational level, No. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>8 (53.3)</td>
<td>7 (41.2)</td>
<td>6 (37.5)</td>
<td>21 (43.8)</td>
</tr>
<tr>
<td>Some primary</td>
<td>5 (33.3)</td>
<td>7 (41.2)</td>
<td>8 (50)</td>
<td>20 (41.7)</td>
</tr>
<tr>
<td>Completed primary or some secondary</td>
<td>1 (6.7)</td>
<td>1 (5.9)</td>
<td>1 (6.3)</td>
<td>3 (6.2)</td>
</tr>
<tr>
<td>Completed secondary or higher</td>
<td>1 (6.7)</td>
<td>2 (11.8)</td>
<td>1 (6.3)</td>
<td>4 (8.3)</td>
</tr>
<tr>
<td>Household food security,a No. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secure</td>
<td>5 (33.3)</td>
<td>5 (29.4)</td>
<td>9 (56.3)</td>
<td>19 (39.6)</td>
</tr>
<tr>
<td>Mildly insecure</td>
<td>2 (13.3)</td>
<td>2 (11.8)</td>
<td>3 (18.8)</td>
<td>7 (14.6)</td>
</tr>
<tr>
<td>Moderately insecure</td>
<td>5 (33.3)</td>
<td>5 (29.4)</td>
<td>3 (18.8)</td>
<td>13 (27.1)</td>
</tr>
<tr>
<td>Severely insecure</td>
<td>3 (20)</td>
<td>5 (29.4)</td>
<td>1 (6.3)</td>
<td>9 (18.8)</td>
</tr>
</tbody>
</table>


### Barriers and Strategies for Improving Adherence

Qualitative analysis of reported adherence barriers and facilitators did not identify notable differences across women with high and low adherence, although there were some differences by assigned regimen. Quotes illustrating the barriers and facilitators discussed below are included in Table 4.

Forgetting supplements when busy with work, chores, or children, or when away from home was the most common barrier to adherence, reported by one-third of the women. The midday tablet in the 4- and 3-event regimens was the most difficult to take consistently, particularly on days with extra activities (e.g., church, market, holidays, weddings) and for women who were responsible for livestock or harvesting. Forgetting decreased over time, and women reported that habituation to the regimen helped them to adhere. As the study progressed, the harvest season ended and more women were at home, which helped adherence, particularly the midday dose.

Most women who reported missing doses when outside the home had thought they would...
### TABLE 4. Quotations Illustrating the Range of Perceived Facilitators and Challenges for Adherence to Calcium Supplementation Across High and Low Adherers, Ethiopia (N=48)^

<table>
<thead>
<tr>
<th>Facilitators</th>
<th>Illustrative Quotations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reminders</td>
<td>I used the calendar... I past it on the wall and always when I see it, I remember to take my pills. —25-year-old woman, 4-event, very high adherence</td>
</tr>
<tr>
<td></td>
<td>My husband goes to work during my pill-taking times, which helps me remember. I also associate taking pills with mealtimes. —22-year-old woman, 3-event, increased adherence</td>
</tr>
<tr>
<td>Value of prevention</td>
<td>I often think, “If I die from some problem, what would happen to my children?” I took these pills as you told me to prevent such risks. —27-year-old woman, 3-event, high adherence</td>
</tr>
<tr>
<td></td>
<td>Other people may say, “Why do I bother to take the pills if I am healthy?” Those who are sick are more concerned with taking pills, but I keep taking them to help my baby survive. —38-year-old woman, 3-event, very high adherence</td>
</tr>
<tr>
<td>No adverse effects and relief of symptoms</td>
<td>Feeling healthy and comfortable helped me take [calcium]. If I were not happy in taking the pill, I couldn’t take it. I may avoid it even. —37-year-old woman, 3-event, very high adherence</td>
</tr>
<tr>
<td></td>
<td>Before I started to take these pills I had stomachaches and back pain. Now I am free from these problems. I like these pills very much due to the benefits I received. —30-year-old woman, 2-event, high adherence</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>My brothers asked me whether [calcium] is for [HIV/AIDS]. I don’t want to hear what other people think. I will use it without fear; it is about my health, not others. —30-year-old woman, 3-event, high adherence</td>
</tr>
<tr>
<td>Travel strategies</td>
<td>I took one [calcium] pill and covered it with clean paper to bring with me. To take one pill, I do not want to carry all the pills in the bottle. If I carry the bottle, people may see it... attracting their attention. —22-year-old woman, 4-event, very high adherence</td>
</tr>
<tr>
<td>Trust in government and health services</td>
<td>I am taking the pills consistently. People talk about the pills fattening the fetus... or other unjustified ideas... but health workers and government give this service for our benefit. —20-year-old woman, 2-event, very high adherence</td>
</tr>
<tr>
<td></td>
<td>Since the government does not give us anything which hurts us, taking these pills is not difficult. —30-year-old woman, 3-event, adherence decreases</td>
</tr>
<tr>
<td>Counseling and follow-up support</td>
<td>Your advice has also greatly influenced me to take the pills because I am convinced about their importance and I was committed to take the pills myself. —20-year-old woman, 3-event, high adherence</td>
</tr>
<tr>
<td></td>
<td>Had I not received your advice on how to take [calcium] I could have stopped taking it. And I would have thrown away [IFA] because of the severe heartburn. After your advice to take it right before sleeping, there is no problem for me. —30-year-old woman, 4-event, very high adherence</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Challenges</th>
<th>Illustrative Quotations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forgetting</td>
<td>It was difficult to take [calcium] every day because we are busy with harvest these weeks and I forget to take them. —18-year-old woman, 3-event, high adherence</td>
</tr>
<tr>
<td></td>
<td>Taking [calcium] in the morning is difficult—breakfast may even be delayed—because when I wake up I directly deal with work and may forget it. —38-year-old woman, 2-event, decreased adherence</td>
</tr>
<tr>
<td>Away from home during the day</td>
<td>Taking [calcium] midday is difficult. I went to visit my parents. I left in the morning but failed to come back the same day. I did not take pills for two days. To overcome this, I have to take the pills with me. —24-year-old woman, 4-event, very low adherence</td>
</tr>
<tr>
<td></td>
<td>Midday, most times I am not at home. I may visit family whose relative died, attend a wedding, or travel elsewhere. In the morning no one leaves without breakfast and in the evening it is a must to come home for dinner, but it is unsuitable to take pills at lunchtime. —28-year-old woman, 3-event, very high adherence</td>
</tr>
<tr>
<td>Stigma and community perceptions</td>
<td>People do not know [calcium]. They say, “During our pregnancy we took [IFA] but we don’t know this one.” People may assume this pill is given for [HIV/AIDS]. This pill is difficult for me because I don’t think people’s presumption will change and I don’t know what to do. —25-year-old woman, 4-event, very low adherence</td>
</tr>
<tr>
<td></td>
<td>Taking these pills is very difficult. People who have seen me taking pills for a long time may suspect I am taking [HIV/AIDS] drugs. I can convince he who dares to ask me, but he who doesn’t know me and doesn’t ask, I don’t have the chance to explain the reality. —37-year-old woman, 4-event, very high adherence</td>
</tr>
</tbody>
</table>
TABLE 4. Continued

<table>
<thead>
<tr>
<th>Challenges</th>
<th>Illustrative Quotations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Side effects</td>
<td>I didn’t take [calcium] for the last week. I frequently felt nauseous and abdominal discomfort so I stopped. I feel real discomfort in taking it. I don’t feel healthy. — 25-year-old woman, 4-event, very low adherence</td>
</tr>
<tr>
<td></td>
<td>When I take [calcium] I tend to vomit. It nauseates me. I drink a lot of water and then swallow it. Without water it would have all come back up. I really took this pill with great difficulty . . . I took only for the sake of its benefit. — 20-year-old woman, 3-event, high adherence</td>
</tr>
<tr>
<td>Lack of food</td>
<td>Most of the time I can’t find food in the morning. If I find a little food, I have to give priority to my kids. I do not want to take [calcium] if I am unable to find at least one mouthful. You told me to take it if I am unable to find food, but I’m afraid of taking pills like that. — 33-year-old woman, 4-event, decreased adherence</td>
</tr>
<tr>
<td></td>
<td>I haven’t stopped taking the pills. Even though I couldn’t get adequate and quality food, I took it with whatever food was available. — 24-year-old woman, 2-event, very high adherence</td>
</tr>
<tr>
<td>Illness</td>
<td>I was sick. I couldn’t eat and I even drank coffee with difficulty. Now I have my health back. I can eat and do my work well. This is why I remember to take the pills. — 27-year-old woman, 4-event, increased adherence</td>
</tr>
<tr>
<td>Large baby</td>
<td>People in our community say during pregnancy pills fatten the fetus. This was discouraging me, but I think its advantage outweighs the disadvantage and I decided to take the pills. — 30-year-old woman, 2-event, very low adherence</td>
</tr>
<tr>
<td></td>
<td>My worry is the pills may fatten the fetus and cause problems during delivery. People in our community said this . . . I think it is [calcium], because it has milk content, but people say [IFA] can also fatten the fetus . . . I was taking it expecting its benefit outweighs this problem. I was waiting for you to clarify the reality. — 28-year-old woman, 4-event, very high adherence</td>
</tr>
</tbody>
</table>

Abbreviation: IFA, iron-folic acid.

* Participants categorized by mean adherence rate to assigned regimen: over one-third (20/48) were “very high” adherers (90%–100%), 9 were “high” adherers (75%–89%), 4 were “low” adherers (45%–74%), and 3 were “very low” adherers (<45%). Twelve women did not fit these categories due to substantial changes in adherence over time: for 7 women adherence was categorized as “increased” and for 5 women adherence “decreased.”

Women acknowledged that lack of community awareness and experience with calcium caused misperceptions. A third of participants experienced side effects that they attributed to calcium (nausea, light-headedness, vomiting, abdominal distension, and constipation) and almost half experienced side effects they attributed to IFA (heartburn, nausea, gastritis, abdominal cramp/distension, headache, and weakness/dizziness). Seven women skipped calcium doses due to side effects; 5 discontinued use. Nine women missed IFA doses; 6 discontinued use. Skipping doses due to calcium side effects declined for 60% of these women by visit 4.

Several women mitigated side effects or unpleasant taste or smell by consuming more water and food with supplements, and some took IFA right before bed or ate something sweet, such as sugarcane, afterwards to reduce the bitter taste and nausea. Two women said that taking calcium concurrently with IFA mitigated IFA side effects. Four women reported that side effects went away on their own. Another 9 women continued supplementation with some discomfort, persevering to obtain the health benefits. Some women noted return in time; they suggested they could solve this problem by carrying the bottle, although most did not report trying this. One-fourth of the women mentioned not taking calcium supplements in public and concealing calcium to avoid questions about the supplements or because they had been discouraged from taking calcium by others who viewed it as HIV/AIDS medication. Five women reported that this possibility affected their adherence; several others did not care what others thought. Women acknowledged that lack of community awareness and experience with calcium caused misperceptions. A few women said that taking supplements in public was inappropriate, particularly during church, burial ceremonies, or other formal events, or during fasting periods, causing them to occasionally miss doses. Women were counseled to take only 1 calcium tablet when opening the bottle so MEMS could successfully record consumption. However, women who took calcium out of the bottle to carry for later usage were less likely to miss midday doses.
they took calcium with ease and no adverse effects.

Counseling materials recommended taking supplements with food and water to minimize side effects, but lack of food and water at dosing times made this difficult for more than 12 women, affecting adherence to calcium or IFA. This difficulty was especially true for women reporting long work hours outside the home or food insecurity. Women reported reserving a small amount of food (e.g., 1 piece of bread) to consume with supplements. One woman requested small amounts of food from her neighbors to take the supplements during a food shortage.

Six women reported skipping calcium and IFA supplements due to illness (kidney problems, gastritis, cough, and weakness); 3 were prescribed medications by a health worker and felt it was too difficult to continue supplementation while taking medication. Adherence usually improved when women regained their health.

When asked if anyone discouraged supplement adherence, many women said family members or others linked supplements with large fetuses and difficult deliveries. Most women felt this discouragement was “unjustified” or “lacked evidence” but 6 were concerned it might be true.

**Facilitators to Adherence**

When asked what helped women remember their tablets, three-fourths of the respondents said reminders from family members and the reminder calendar, including encouragement from others to use the calendar or husbands or children marking the calendar when supplements were taken. A few women used the calendar to explain supplementation to others. Keeping pills visible and taking supplements during routine events (e.g., mealtimes or coffee breaks) also helped many women adhere.

Most women were motivated by perceived health benefits from calcium or IFA supplements, especially relief from light-headedness, dizziness, and fetal movement. Over time, more women reported that anticipated and experienced health benefits facilitated their adherence. Over a third of the women said they were motivated to take calcium and IFA to supplement what they viewed as their own poor diets, restating counseling messages used by interviewers that tablets provided nutrients that otherwise needed to be obtained from foods they could not regularly access (e.g., milk, liver, or meat).

Over half of the women mentioned trusting that health workers and the government would only provide tablets that were good for them. This trust did not correlate with adherence, due to the influence of other barriers (e.g., forgetting, side effects), but strongly motivated overall acceptance of supplementation in this context. Similarly, many women noted that the counseling and advice received from health workers and the study interviewers encouraged them to continue taking the supplements, especially when they had doubts or challenges.

**Calcium Product Preference and Acceptability**

At study enrollment, about three-fourths of the women selected chewable supplements for their sweet taste and smell, and sometimes for ease of consuming without water and a belief that conventional tablets are difficult to swallow. Women who initially chose conventional supplements cited simplicity of use, absence of strong taste, smaller size, and concern that chewable supplements could cause nausea.

Over half of the women maintained their original choice throughout the study, citing satisfaction with supplement properties, absence of challenges or side effects, familiarity, and dislike or uncertainty about the other supplement type. However, about a third of the women who selected chewable supplements switched to conventional due to untoward effects (e.g., excessively sweet taste and foamy mouth) or side effects (nausea, vomiting, and heartburn). Four women who changed from conventional to chewable supplements cited lack of access to water; 1 woman disliked the taste. Most who switched supplement type reported full correction from previous challenges. At final interviews, among the participants willing to continue calcium until delivery, about 60% chose chewable supplements.

**Regimen Preference**

Although regimens were originally randomly assigned, at visit 3 participants were shown illustrations of 3 dosing regimens and asked to choose a regimen for the final 2 weeks. The 2-event was preferred; nearly all participants assigned this regimen (88%; 14/16) chose to continue at visit 3 compared with fewer than half (41%; 7/17) of the women assigned to the 3-event and very few (27%; 4/15) assigned to the 4-event. Overall, two-thirds of the women (20/32) switched to the 2-event versus 2 women switching to the 3-event
Women chose the regimen with the lower calcium dosage due to difficulties adhering to the midday supplement.

**Calcium Adherence and Consumption**

MEMS data indicated that most women (42/48) initiated calcium supplementation immediately, and all women eventually tried supplementation. Across the sample, mean adherence rates after 4 weeks were high (approximately 80% for all regimen groups), with no associations between adherence rate and assigned or chosen regimen (Table 5). Throughout the study, most women (35/48) maintained average adherence rates of over 80% or increased their adherence to this level; a few (7/48) completed the study with a mean adherence rate of less than 50%.

A one-way ANOVA showed a significant difference in the mean calcium consumption between randomly assigned regimen groups, but not in adherence rate. Post hoc comparisons using the Tukey HSD test (Table 5) indicated that women assigned the 2-event regimen on average consumed significantly less calcium (811.3 ± 227 mg/d) than women assigned the 3-event (1,251.1 ± 211.9 mg/d) or 4-event (1,156.4 ± 514.5 mg/d). Women assigned the 3-event regimen consumed the most calcium because of the 1,500-mg dose combined with the highest adherence rate, but the trend was not statistically significant. In the later phase in which women chose their regimen, results were similar; calcium consumption was lower among women who chose the 2-event regimen and adherence rates did not differ. Statistical results for this later phase exploring women’s choices must be interpreted with caution given the wide range in group sizes due to the self-selection that was inherent in giving women a choice.

Although the 2-event regimen was preferred, switching to this regimen did not improve the adherence rate, so mean daily calcium consumption for the 20 women making this selection decreased from 1,134.6 mg (±454.4) to 752.8 mg (±327.8). Table 5 shows the trends in mean daily calcium consumption between women’s assigned and chosen regimens. Adherence was generally high, thus most women consumed on average at least 800 mg/d regardless of regimen type. Poor adherers tended to choose the 2-event regimen yet continued to struggle with adherence, yielding < 500 mg daily consumption for 25% of this subgroup. While some women reported that aspects of regimens were difficult to adhere to, having a choice of regimen did not substantially affect adherence. All women who were assigned or chose the 3-event regimen consumed on average at least 500 mg/d; most consumed at least 800 mg. Women selecting 3- and 4-event regimens were already highly adherent with no notable side effects and believed the new regimen with additional calcium would increase health benefits.

At study end, almost all women were willing to take calcium and IFA until delivery and three-fourths chose the 2-event regimen. Five women.

**TABLE 5.** One-Way ANOVA for Antenatal Calcium Supplementation Adherence Rate and Daily Calcium Consumption, by Randomly Assigned Regimen (for 4 Weeks) and Chosen Regimen (for 2 Weeks) Among Ethiopian Women in 2 Rural Districts (N=48)²

| Assigned regimen (weeks 0–4) | 4-Event | 3-Event | 2-Event | F₂,₄₅ | P Value §
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Adherence rate, mean % (SD)</td>
<td>n=15</td>
<td>n=17</td>
<td>n=16</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>77.1a²</td>
<td>83.4a²</td>
<td>81.1a²</td>
<td>.265</td>
<td>.768</td>
</tr>
<tr>
<td></td>
<td>(34.3)</td>
<td>(14.1)</td>
<td>(22.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calcium consumption, mg/d, mean (SD)</td>
<td>n=15</td>
<td>n=17</td>
<td>n=16</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1,566.4a²</td>
<td>1,251.1a²</td>
<td>811.3b²</td>
<td>7.53</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>(514.5)</td>
<td>(211.9)</td>
<td>(227)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chosen regimen (weeks 4–6)</td>
<td>n=5</td>
<td>n=9</td>
<td>n=34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adherence rate, mean % (SD)</td>
<td>n=5</td>
<td>n=9</td>
<td>n=34</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>95.9a²</td>
<td>80.1a²</td>
<td>73.8a²</td>
<td>1.248</td>
<td>.297</td>
</tr>
<tr>
<td></td>
<td>(3.9)</td>
<td>(23.1)</td>
<td>(32.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calcium consumption, mg/d, mean (SD)</td>
<td>n=5</td>
<td>n=9</td>
<td>n=34</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1,398.5a²</td>
<td>1,200.9a²</td>
<td>738b²</td>
<td>15.66</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>(58.3)</td>
<td>(347.1)</td>
<td>(328.3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: ANOVA, analysis of variance.
² Groups followed by the same letter are not significantly different in post hoc multiple comparisons with Tukey correction, P<.05.
§ P value for overall ANOVA F-test.
TABLE 6. Trends in Average Daily Calcium Intake at 4 Weeks (Assigned Regimen) and 6 Weeks (Chosen Regimen) Within Household Trials Among Participating Ethiopian Women in 2 Rural Districts (N=48)\(^a\)

<table>
<thead>
<tr>
<th>Calcium consumed, mg/d(^b)</th>
<th>Assigned regimen (weeks 0-4)</th>
<th>Chosen regimen (weeks 4-6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4-event (n=15),(^b) %</td>
<td>3-event (n=17), %</td>
</tr>
<tr>
<td>&lt;500</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>500 to &lt;800</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>800 to &lt;1,000</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>1,000 to &lt;1,500</td>
<td>47</td>
<td>82</td>
</tr>
<tr>
<td>1,500</td>
<td>27</td>
<td>6</td>
</tr>
</tbody>
</table>

\(^a\) Cutpoints based on 500 mg calcium per supplement.  
\(^b\) Because of rounding, percentages may not total 100.  
\(^c\) Estimated average requirement (800 mg/day) and recommended dietary allowance (1000 mg/day) for pregnant women aged 19–50 years based on Dietary Reference Intakes for Calcium and Vitamin D.  

were unwilling to continue calcium and 4 were unwilling to continue IFA. Among women discontinuing calcium, 3 had low mean adherence throughout the study (12%–46%); 1 woman was 100% adherent but had severe heartburn and nausea; and 1 woman (87% adherent) felt she had taken enough.

**DISCUSSION**

Hypertensive disorders in pregnancy are a leading health problem in Ethiopia, and early screening and prevention strategies are needed.\(^{51,52}\) In response to WHO guidelines,\(^7\) our study explored modifiable factors associated with adherence to calcium supplementation and identified strategies to optimize implementation of calcium with IFA supplementation.

Women in Oromia, Ethiopia, found antenatal calcium and IFA supplementation acceptable. All participants were willing to try supplementation and most adhered to calcium supplements at relatively high rates over a 6-week period, after receiving supplements and behavior change communication to motivate and support adherence. Perceptions of physical benefits and the need to supplement poor diets were motivating, and social support and reminder materials helped to overcome barriers and enhance adherence.

A comparison of our results with parallel research in Kenya\(^{53}\) identified important similarities and differences across settings. Calcium consumption and adherence rates in Kenya were generally lower. Most participants in Ethiopia and Kenya (~75%) preferred chewable supplements; however, in Ethiopia, the difference in preference at the end of the trial was negligible. Results in Ethiopia indicate that a sweet chewable tablet did not strongly influence adherence and, given the significantly higher cost, inclusion of this option in programs is not warranted. However, preferences are likely to vary by context.\(^{23,53}\)

Side effects of calcium or IFA were widely reported, yet only affected adherence for a subgroup of the women, in contrast to the reported impact of side effects (or fear of them) on IFA adherence in previous Ethiopian studies.\(^{54,55}\)

Counseling and regular follow-up visits as part of the TIPs approach may have positively influenced women’s adherence, especially given high levels of trust in medical providers in Ethiopia.\(^17\)

Women who know the importance of IFA are more likely to adhere\(^{21,54,56}\); however, our results and previous formative research indicate that follow-up counseling to monitor adherence and address challenges such as side effects is as important as initial education.\(^21\) Utilizing extension officers as agents for sharing information on antenatal supplements has been recommended,\(^{21}\) and continued support of the program could leverage their home visits and trusted relationships to follow-up on supplementation.\(^{37}\) As in formative research,\(^21\) counseling on supplements for prevention and as a dietary
Strategies for improving adherence were discussed during counseling and were more effective than reduced dosage for boosting adherence.

The Ethiopian government has highlighted prenatal calcium supplementation in national health policy, but it has yet to be implemented as a part of ANC.61 Our research suggests the importance of targeting communities, not just pregnant women and their families, when expanding antenatal supplementation programs. Over 40% of the sample mentioned stigma due to pills being associated with HIV/AIDS or concern that supplements cause excessive fetal growth and difficult deliveries, beliefs noted previously in Ethiopia.17,21,55 Discouragement often came from individuals not typically targeted for antenatal supplementation counseling, including extended family members, neighbors, and other community members. Broad community-level sensitization on calcium and IFA supplementation could help normalize taking multiple tablets during pregnancy as a preventive measure, particularly where preeclampsia is not well known.53 Community-based mobilization to promote IFA supplementation has been successful elsewhere.62

The level of calcium supplementation must be considered in relation to dietary calcium intake during pregnancy. In Ethiopia, dietary calcium intake during pregnancy is often low.63 However, dietary intakes are variable such that the preferred 2-event regimen could be sufficient for some women to meet requirements, and higher doses could exceed the tolerable upper intake level.64 Programs that strengthen dietary quality and diversity should also be considered and, if feasible, may be more sustainable. Consumption of teff, a common local grain,65 and diverse diets66,67 have been associated with anemia reduction in Ethiopia and dietary diversity lowered preeclampsia risk in other populations.68,69 Diversifying diet in addition to IFA supplementation was associated with reduced occurrence of symptoms suggestive of preeclampsia and eclampsia in India.70

Strengths and Limitations

The TIPs approach gathers data on emic views and actual experience with antenatal supplements and can provide effective guidance to strengthen nutrient programs.43,53,60 However, home delivery of supplements, repeated counseling during household visits, and purposeful sampling limits generalizability of the findings. This intensive model facilitated in-depth exploration of the complex factors affecting acceptability and adherence, but facility-level implementation research is needed to determine feasibility and sustainability through the health system in Ethiopia and in other contexts.
Researchers undertook a larger, facility-based study in Kenya in order to further explore findings from TIPS.71 Use of electronic monitoring in combination with interview data to assess participants’ adherence strengthened study conclusions. Agreement between MEMS and self-report data varied across the sample, so our report focused on MEMS adherence data, which appeared to be more accurate. For example, most women whose adherence (measured by MEMS) was categorized as low or very low or whose adherence decreased throughout the study regularly stated that they never or only rarely missed a supplement and that taking calcium was “easy.” MEMS allowed for cross-checking against this self-report data. Upon further probing, 8 women reported different (and multiple) barriers. Consistent overreporting of product use is well documented72,73 and limits researchers’ ability to accurately measure adherence.

MEMS also posed challenges. Despite interviewer instruction, some women initially struggled to open the MEMS cap resulting in husbands helping with “practice openings.” For this reason, electronic monitoring data were capped at 100%. MEMS is limited as an effective monitoring tool for dosing regimens that include separate pills for more than one micronutrient. Thus participants were provided blister-packaged IFA separately, and IFA adherence was not electronically monitored.

Women’s reluctance to take calcium in public may have been exacerbated by the large size of the MEMS cap and instructions to travel with the bottle rather than remove tablets for later dosing. “Pocket dosing” facilitated taking a midday dose but could have resulted in inaccuracies in MEMS data. Additionally, knowing that pill-taking was monitored may have increased adherence for some.

Although regimens were assigned randomly, the group assigned the regimen with only 2 pill-taking events appeared to have lower rates of food insecurity. One would expect that if this had any effect, it would be to increase adherence, given that food insecurity was mentioned as a barrier, but we did not find higher adherence in this group. Thus, the variation in food security does not change our conclusion that reduced pill-taking events did not result in higher adherence.

Finally, this study was conducted in an area recently served by a maternal health intervention that strengthened various services related to pregnancy care and IFA supplementation41; motivations, preferences, and adherence to supplementation in other less well-established program areas may reveal different results.

[CONCLUSION]

To reduce preeclampsia and anemia, the WHO recommends antenatal calcium and IFA supplementation in populations with low dietary intake. Concurrent antenatal calcium with IFA supplementation was acceptable for most women in this setting, even in the context of negative community perceptions of pill-taking and low IFA use before the study. Women were eager for extensive information about supplementation and motivated by knowledge about health benefits of calcium and IFA. Social support, regular counseling addressing key barriers, and home-based reminders facilitated calcium adherence. Due to the complexity of the prenatal calcium regimen in the WHO guidelines (1.5 g/d in 3 doses with separate IFA administration), we compared adherence to simpler regimens. Calcium adherence did not differ across regimen; thus, a regimen of fewer calcium doses (1 g/d) that women preferred resulted in lower calcium intakes. A regimen of 3 doses of calcium (1.5 g/d) simplified to allow co-consumption with IFA resulted in over 80% of the women consuming 1,000 mg or more daily. This approach is recommended, pending further research to identify the minimal effective dose to reduce preeclampsia risk. Determining the feasibility of integrating calcium into existing antenatal IFA supplementation programs and achieving high levels of adherence requires further implementation research within health systems, but results indicate high acceptance of calcium supplementation in this context.

Acknowledgments: We are grateful to the participants and local health staff who shared their valuable time and experiences with us. We thank the data collection team, including Mamusha Aman, Keneni Gutema, Abebe Mamo, Yohannes Mulugeta, and Gemedo Zeleke. We acknowledge technical input and guidance from Rebecca Stoltzfus, Lainre Omotayo, Binoyam Tesfaye, and Crispin Ndedda. Sarah Luna and Salma Mutwafy provided quantitative and qualitative data analysis support.

Funding: This research was supported by Global Affairs Canada through a grant given to Nutrition International (formerly Micronutrient Initiative).

Competing interests: None declared.

REFERENCES


2. Casey GJ, Sartori D, Horton SE, et al. Weekly iron-folic acid supplementation with regular deworming is cost-effective in preventing...


35. Omotayo MO, Dickin KL, O’Brien KO, Neufeld UM, De-Regil LM, Stoltzfus RJ. Calcium supplementation to prevent preeclampsia:


69. Lango-Mbenza B, Tshimanga BK, Buassa-bu-Tsamba B, Kabangu JR. Diets rich in vegetables and physical activity are associated with a decreased risk of pregnancy induced hypertension among rural


What Makes a National Pharmaceutical Track and Trace System Succeed? Lessons From Turkey

Koray Parmaksiz, Elizabeth Pisani, Maarten Olivier Kok

Key Findings

- Turkey implemented the world's first full pharmaceutical track and trace system.
- Its success depended on 4 factors:
  - Political determination
  - Industry incentives
  - Reimbursement linked to verified dispensing
  - Flexible implementation

Key Implications

- Policy makers should perform a systemic analysis of market, political, economic, technical, and legal factors before implementing any pharmaceutical track and trace system.
- To achieve the intended outcomes, system design must align with the goals of implementation (e.g., to tackle fraud, reduce falsified medicine, minimize shortages). The system must be feasible in the local political and economic context.

ABSTRACT

Background: Track and trace systems are increasingly being implemented as a technological solution to secure pharmaceutical supply chains. Turkey was the first country to implement a full pharmaceutical track and trace system throughout the entire regulated domestic supply chain. This article explores the emergence and functioning of this system and the consequences for substandard and falsified medicine with a focus on the underlying political and economic factors.

Methods: This study uses an explanatory case study approach that combined interviews with purposefully selected key informants and document analyses.

Results: The main drivers for implementing the pharmaceutical track and trace system in Turkey centered on the elimination of reimbursement fraud and the prevention of falsified medicine in the regulated supply chain. Although stakeholders experienced both physical and software-related problems in implementation, the alignment of incentives of all stakeholders with the power of the state, along with leeway for adaptations, ultimately resulted in a successful process. This track and trace system provides a clean regulated supply chain, minimizes reimbursement fraud, facilitates fast market recalls, and can flag likely medicine shortages. Staff previously engaged in pharmacy inspections now concentrate on ensuring production quality, which reduces the risk of substandard medicines.

Conclusions: In Turkey, 4 factors drove the successful implementation of pharmaceutical track and trace: the political determination to eliminate reimbursement fraud, a large pharmaceutical market dominated by a single payer, medicine reimbursement being contingent on verified dispensing and prescription, and flexibility to adapt the system according to the needs of stakeholders during implementation.

INTRODUCTION

Track and trace systems are logistical technologies that enable localizing and following a product throughout a supply chain; they are used in many sectors, including aviation and retailing. In 2012, Turkey became the first country in the world to implement a full track and trace system to secure its domestic pharmaceutical supply chain. A growing number of countries are now following suit. Argentina and Saudi Arabia are among the countries that have already put such a system in place, while other countries, including China, the United States, and European Union (EU) member states, are currently in the process of implementation.
For pharmaceuticals, 2 track and trace systems dominate. The first, known as a “point-of-dispense check” system, validates medicine packages at the points where they are dispensed to patients (e.g., pharmacy or hospital) with the code assigned during the manufacturing process. Other transactions (e.g., between wholesalers and distributors) are not systematically recorded. The European Medicines Verification System, implemented by the EU in February 2019 as part of the Falsified Medicines Directive, provides an example. This system allows for verifying the authenticity of the product, but not for tracking the product throughout the supply chain. The second track and trace system, often referred to as full track and trace, validates a medicine package at every stage of its journey through the supply chain. This system is in use in Turkey.

Although the full track and trace system is more complex to implement, it provides additional potential benefits to those of the point-of-dispense check system. These benefits include real-time tracking throughout the entire supply chain, stock management for timely detection and prevention of stock-outs, targeted product recalls, and reduction of reimbursement fraud, theft, and medication errors.

Recent studies show that the opportunity to enter the pharmaceutical supply chain differs between substandard and falsified medicines. Substandard medicines, which are defined as authorized products that fail to meet quality standards, enter the supply chain through manufacturers who might sacrifice quality to maximize profits. Falsified medicines, which have a deliberately misrepresented identity, composition, or source, are often introduced by criminals who see a market opportunity when shortages of quality and affordable medicine occur in the regulated market.

Turkey began with a point-of-dispense check system known as Ilaç Takip Sistemi (ITS) in 2010, before introducing full track and trace in 2012. From the start, all medicines sold in Turkey had to be equipped with a DataMatrix code, which is a 2-dimensional barcode. A DataMatrix code contains information on the Global Trade Item Number, a serial number, an expiration date, and a batch number, which enables tracking the history and location of each medicine through the supply chain.

Implementing New Technologies

In this case study, we draw upon insights into the implementation of new technologies, which are based upon the rich literature on technological innovation. First, technological changes do not take place in a vacuum; they are embedded in dynamic and complex systems and are shaped by social, political, and economic factors. Therefore, when analyzing the implementation and functioning of new technologies, one must consider the context in which they are embedded.

Second, initiation or successful implementation of technological change is often contingent on the determination of key actors to solve a perceived problem. The process of problematization and the emergence of a relative consensus about the nature of a problem are important first steps in aligning the incentives of all parties playing a key role in implementing the new technology.

Third, implementation and innovation are intimately and reciprocally connected, which means that innovations transform during implementation. As a result, the technology that ultimately gets implemented often deviates from the initial plan. These adaptations, which are often omitted in retrospective accounts about the success of a technology, enable a certain system or technology to operate successfully within its specific context. Overlooking these adaptations undermines the successful reproducibility of technological innovations. In addition, adaptations enable different actors to assign different roles to the technology that are not limited to the official function of the technology. This results in unforeseen outcomes of the technology that are valuable to capture in order to reach its full potential.

Although Turkey was the first country to implement a full track and trace system, neither how the country achieved this significant feat nor what made it possible has been closely investigated. The aim of this study was to gain insight into political and economic factors that drove the implementation of the pharmaceutical track and trace system in Turkey. We paid special attention to the consequences of the system for substandard and falsified medicines. Insights from our study may provide valuable knowledge to other countries aiming to implement pharmaceutical track and trace systems and may contribute to the understanding of implementing large technological systems in the health sector.

**METHODS**

**Study Design**

For this qualitative case study, document analysis and semistructured interviews were carried out. We used an explanatory case study approach, the main purpose of which was “to explain how and why some conditions came to be.” Such an
approach allows for investigating underlying factors that are often too complex to be captured by surveys or other quantitative measures alone. We believe that a detailed understanding of the implementation and adaptation of the pharmaceutical track and trace system will provide valuable insights into the complexities involved in implementing such large-scale health technologies.

Study Participants
The study participants were 16 purposefully selected key informants. Selection was based on their knowledge and expertise in political and economic factors influencing the emergence, implementation, and functioning of ITS. The aim of purposeful sampling is to increase depth and richness of the collected data by identifying and selecting information-rich cases from different perspectives. We sought to involve stakeholders across the supply chain, together with independent experts, to achieve a comprehensive evaluation of the pharmaceutical track and trace system in Turkey. Backgrounds of study participants are shown in the Table.

Data Collection Methods
To prepare for interviews and to triangulate findings, we reviewed relevant policy documents that helped us understand the emergence, implementation, and functioning of ITS. This method involves an iterative process, in which newly collected data in the form of interviews were triangulated with existing data from previous interviews, studies, or reports obtained during literature research to inform subsequent data collection and verify findings. First, the semistructured interviews were recorded and transcribed verbatim. Then, if necessary, they were translated into English. Interviews were coded using a coding structure jointly developed by the research team. This structure was based on political and economic factors enabling market opportunities for substandard and falsified medicine, along with factors facilitating or obstructing the implementation and functioning of track and trace systems. Emerging themes and the analysis were discussed during 6 weekly team meetings, until consensus was reached. NVivo (12.0.0) was used as the qualitative data analysis software.

RESULTS

Historical Developments and Pricing Policies
To provide the contextual background, we asked participants to reflect on the historical developments of the health sector and the pharmaceutical industry in Turkey. Until the early 2000s, Turkey experienced several problems in the health sector, including insufficient insurance coverage, poor health outcomes such as life expectancy and maternal mortality, and relatively low governmental health expenditure.

In addition, Turkey had 3 state institutions, which are known by their Turkish abbreviations SSK, BAĞ-KUR, and Emekli Sandığı, providing health insurance to different employment-based groups. These institutions operated independently, which resulted in high fragmentation of service provision and restricted access to health services.
At that time about half of Turkey’s population, 50 percent was covered under SSK, and the number of hospitals those people could use was only 120. For the whole of Turkey, can you imagine? Half of the population in Turkey is doomed to only 120 hospitals. —Academic

After the national elections in 2002, the Justice and Development Party came into power in Turkey. This was the first time a political party with religious roots came into power as a single-party government since the establishment of the constitutionally secular Republic of Turkey. Therefore, they had to establish political legitimacy among their citizens and the international community. The government made a strong commitment to universal health coverage as a way of establishing political legitimacy among the country’s citizens and in the international community. The government subsequently introduced the Health Transformation Program in 2003 with the aim to increase insurance coverage and financial risk protection.

After 2006, the government merged the 3 state institutions providing health insurance to form a single-payer institution, called Sosyal Güvenlik Kurumu (SGK).30 Insurance coverage provided by SGK increased access to health services in Turkey considerably, resulting in a significant increase in public health expenditure. In response, the government introduced price-cutting measures, such as reference pricing in 2004 and global budgeting between 2010 and 2012. Despite these measures, manufacturers continued to supply the Turkish market, mainly because increased access to health services increased the overall volume of sales, creating a substantial pharmaceutical market that manufacturers were not willing to give up:

There was not such a thing as convincing. The state is not obliged to convince. The customer is king. “I pay the money; I determine the conditions.” Turkey has such an advantage. I buy more than 80 percent of the market. They say: “If you are willing to give [medicines] under these conditions, then you can give them. Otherwise, I’m sorry, go sell them in another country, don’t sell them to me.” —Multinational manufacturer

Respondents were asked if downward price pressures incentivized manufacturers to cut corners, resulting in substandard medicines. Both manufacturers and the Ministry of Health (MOH) emphasized that the production or import of substandard medicine in the Turkish market was very unlikely because Turkey has well-defined legislation and regulations, including Good Manufacturing Practices (GMP), inspections, laboratories, and a pharmaceutical track and trace system. According to respondents, this strong regulatory framework minimized the possibility of substandard products on the market, while enabling rapid detection.

**Pharmaceutical Track and Trace**

An MOH official explained that before the introduction of the pharmaceutical track and trace system in 2010, quality assurance of medical products was mainly based on market surveillance and inspections. When pharmacists, health professionals, or others reported suspicions about a product, the MOH might sample that product for testing. This largely reactive system was time and resource intensive.

Despite the successful pricing policies to reduce medicine prices, the state experienced significant losses due to fraud around 2007. Although falsification and theft contributed to these losses, all respondents indicated that the main reason for the implementation of the pharmaceutical track and trace system in Turkey (i.e., ITS) was the presence of “reimbursement fraud” or “barcode scamming.” Prior to ITS, pharmacies had to cut out the barcode of each product sold and put it behind the invoice. The invoice would then be sent to SGK for reimbursement. However, this system was vulnerable to fraud, as seen in the following example:

_I know your national identity number. I am a doctor and I am writing the prescription to you, but you don’t know that I am writing it. I give this prescription to the pharmacy. And the pharmacy doesn’t sell the drug but sells the barcode. It prints the barcode on the offset and sticks that barcode behind that prescription or the invoice and sends it to SGK. Takes the money, but there is no transaction or trade. I mean, nobody sells anything, but gets the money from SGK. It is not a fraud to people; it is a fraud to the government._—Technical agency official

Respondents mentioned the existence of “printing houses” exclusively printing these falsified barcodes to be reimbursed by SGK. This fraud was estimated to account for US$1 billion annually.31,32

**Implementation of ITS**

To prevent reimbursement fraud, the government decided to implement ITS. The first discussions on ITS occurred in 2007, but it took 3 years to convince and prepare stakeholders. The implementation took place in 2 phases to reduce implementation problems. A few months prior to the implementation of the first phase, a short pilot
The study was done with a small number of pharmacies. Then, phase 1, which focused on the manufacturers and pharmacists, was introduced in 2010. These 2 stakeholder groups, which are the front- and tail-end of the pharmaceutical supply chain, were obliged to make sales notifications, but wholesalers were not yet included in the system. Respondents mentioned that the phase 1 system, which corresponded to a point-of-dispense check system, remained vulnerable to introduction of falsified medicines at points in the supply chain where transactions were not tracked.

The initial implementation of phase 1 was problematic because of the way in which the software was developed. A single software engineer who had little experience in building enterprise systems was given the responsibility to create the software. The system crashed shortly after phase 1 was launched in 2010 and the system then had to be rebuilt from scratch by another person. Respondents indicated that the political will and determination of senior political figures were the main driving forces for this project to succeed:

The undersecretary general called me to SGK, I was shopping in my sportswear at that time. He told me: “Okay, it is not important, come here!” I was running in the hall of SGK and I opened the door of the meeting room. I was sweating, in sportswear and I saw that two ministers, two undersecretary generals, two vice presidents, a lot of big guys were in the meeting room. I was shocked. The undersecretary general said: “This is the person I mentioned.” I sat between two ministers and they asked: “The system crashed, what should we do?” I told them: “First, accept it. In front of the news, in front of the media, first accept it and postpone it.” They said: “This is politics, we cannot do that. You have 10 days, please make it work.” This is the Eastern culture.—Technical agency official

Phase 2 was implemented in 2012. It can be described as full track and trace, encompassing all actors within the domestic regulated supply chain. It was based on cross-checking movements of a product between each actor by comparing sales and purchase notifications. After phase 2, maintenance and development of ITS came under the responsibility of another company and MOH. The data on the medicines were pooled in a centralized database managed by the MOH. A schematic representation of the ITS workflow is shown in the Figure.

**Reaction to ITS**

When ITS was first introduced, the expectations on its feasibility varied widely between different actors. Government institutions, including MOH and SGK, were convinced that the system would succeed, and that it would cut fraud and thus expenses in the national health system. Manu-

---

**FIGURE.** Simplified Schematic Representation of the Workflow of the Full Pharmaceutical Track and Trace System Used in Turkey
facturers, however, did not think it was feasible. No other country had successfully executed national track and trace, and manufacturers were especially skeptical that it could be achieved in the very limited timeframe envisaged by Turkish politicians:

“There was a thing like: “Well, it won’t be implemented here anyways. It will fail. Let’s act as if we are complying to it, it won’t work anyways.” —Multinational manufacturer

Industry was unhappy about having to bear the costs of compliance. One manufacturer estimated that his company invested around US$5 million in adding track and trace to their production and distribution flow, and that amount did not account for production losses. Another manufacturer said that the costs were around US$100,000 per conveyor belt. An additional concern of manufacturers was the lack of adequate equipment:

“We invited the Germans and Italians, who were good in machines. We talked to them. When we first bought it, we were buying dreams. (...) The ones who were able to convince you the most, you picked, because there was nobody that could show you how it would work). —Multinational manufacturer

The wholesaler, who shared these initial doubts and skepticism, explained that they invested around 100 million Turkish Liras to implement this system across Turkey.

The merger of the 3 state payers into a single health insurer provided the state with consolidated buying power covering around 95% of the market, which was sufficient to incentivize the pharmaceutical industry to implement the system. Manufacturers also saw a benefit in reducing access to the market for medicine falsifiers, thus protecting income and value.

The pharmacists included in our study indicated that they were most concerned about the increased workload because each product had to be individually scanned. Forgetting to do so could have serious consequences for the pharmacists during inspections.

With all these different interests, the challenge of aligning personal and institutional incentives was far greater than any technological challenge, as an MOH official pointed out:

“I can get four people from India who can build this system in a short amount of time. The most crucial part is aligning all stakeholders with the support of the government.” —MOH official

Implementation and Adaptation

Despite the phased implementation, stakeholders experienced many implementation problems, both physical and software related. In phase 1, issues with software development and realistic planning of the implementation process were experienced by manufacturers, wholesalers, and pharmacists. Additionally, a manufacturer explained that physically adapting the production lines to print and scan DataMatrix codes turned out to be challenging:

“We experienced a lot of problems. The ink got wiped because it did not dry properly. Also, when the conveyor belt was a bit skewed, the scanner could not read the DataMatrix code.” —Multinational manufacturer

Existing production lines in factories were not designed to be adapted, while new production lines were not yet developed. In addition, manufacturers experienced problems with sales notifications. Originally, the DataMatrix code on each secondary medicine package—the packaging enclosing the primary packages (e.g., blister or bottle)—had to be scanned individually, which took time and resources. Therefore, manufacturers introduced a “minimum-order-quantity system,” in which wholesalers were obliged to purchase medicines in fixed amounts. This system allowed manufacturers to scan the DataMatrix code on the tertiary packaging (i.e., the shipping-level packaging surrounding the secondary packaging).

Similarly, changes were made to logistic units.25 The industry realized that the serial shipping container code was more appropriate than the serialized global trade item number, which MOH initially proposed. However, respondents pointed out that MOH was willing to modify the system accordingly:

“The things that the authority sees from above and the reality we work in at the operational level are different…. We realized these things by experience. Therefore, things that were written down in theory evolved towards the reality of daily life in the end. Otherwise, if they did not change, if the ministry of health did not take our feedback into consideration, this system would be a non-operative system.” —Multinational manufacturer

The wholesaler experienced similar implementation problems. Scanning the DataMatrix codes and adapting conveyor belts proved challenging. Also, staff needed to be trained to use the system appropriately. Most of the problems experienced by pharmacists related to software malfunction. When the system was inaccessible, pharmacists were not able to sell their products to patients.
problems were more acute in hospitals, where scanners were often in short supply. Additionally, hospitals buying common medicines in bulk found it difficult to scan individual prescriptions because the DataMatrix codes were not affixed to the packages of pills given to the patients but were only on the outer (or secondary) packaging, which was often thrown away before medicines were dispensed as individual patient prescriptions. As a result, hospitals were making consumption notifications instead of sales notifications.26

Although MOH was open and collaborative in adapting the system according to the needs of those involved during implementation, solving practical problems remained largely the responsibility of stakeholders.

Outcomes of ITS
Respondents indicated that implementing ITS has had 5 main positive outcomes. First, reimbursement fraud is highly unlikely to happen in the current system. Successful fraud would require the participation of a long chain of people, raising the risk and reducing the reward for fraudsters. Reimbursement agency officials added that fraud cannot be reduced to zero but only minimized to a certain level, which is believed to have currently been achieved.

Second, ITS has largely eliminated falsified medicines in the regulated domestic supply chain. According to respondents, it is currently close to impossible to sell falsified medicines to patients in pharmacies and get reimbursement from SGK. Medicines cannot be “injected” into the supply chain at any stage other than by the manufacturer or importer. The only possibility for selling falsified products to patients is through out-of-pocket payments. However, since SGK provides comprehensive coverage to almost the entire population, citizens have no incentive to look outside the regulated supply chain or buy products out-of-pocket. Respondents underlined that Turkey’s health financing system, which reimburses pharmacists for verified dispensing and prescriptions, is central to the success of ITS. If pharmacists do not scan the DataMatrix code at dispensing, which serves to verify the authenticity of the product, SGK will not pay them for the product. In the case of a prescription medicine, SGK also verifies the authenticated product with the patient’s medical prescription before paying the pharmacist.

Third, respondents said that ITS has optimized the recall process for products that are degraded or show unwanted side effects. ITS enables quick and targeted recalls of specific products. The sale of suspect products can also be blocked within the system by MOH officials,24 which prevents further dissemination of poor-quality medicines and potentially significantly reduces public health harm caused by them.

Fourth, a mobile application of ITS, which was launched in 2014, allows citizens to scan the DataMatrix code of products. Citizens can immediately check the legitimacy of the product and obtain additional information, such as the expiry date, price, and recall status. Also, side effects can be entered in the application, which facilitates the collection of pharmacovigilance data.

Fifth, since ITS registers sales throughout the supply chain and eventual dispensing by outlet, the system allows for close monitoring of medicine stocks by health authorities, as well as providing inventory control for manufacturers, wholesalers, and pharmacies.

Although ITS has had many positive outcomes, respondents noted that ITS does not guarantee product quality. If a product has poor quality at manufacture or import, it will continue through the supply chain; careful tracking also does not protect against degradation. However, by reducing time spent on pharmacy inspections, ITS allows the transfer of human resources to other quality assurance functions:

Before ITS, we had 3,000 inspectors [checking pharmacies]. After ITS, we have 100 inspectors. The other ones, we didn’t fire them, the other ones are used inside a new department which is GMP compliance, GDP [Good Distribution Practice] compliance and they are taking more samples from the market. They are going to the manufacturing sites and inspecting for substandard products. They are inspecting the active pharmaceutical ingredients. Now, they have time to inspect these things. —Technical agency official

Although quantitative data are not available, respondents reported that the tracking capability of ITS in combination with sufficient and qualified human resources has significantly increased the possibility of detecting substandard and falsified medicine.

Future Adaptations to ITS
Respondents suggested 2 potential improvements to ITS. First, the scope of products given a DataMatrix code should increase. Currently, almost all medicines under the responsibility of MOH are obliged to carry a DataMatrix code, and internet sales are prohibited. However, some...
products, such as intravenous and radiopharmaceutical products, active pharmaceutical ingredients, and personalized medicines compounded in the pharmacy, are excluded from the DataMatrix code requirement. In addition, products such as vitamins and dietary supplements that are overseen by the Ministry of Food, Agriculture, and Livestock are not included in ITS. Respondents pointed out that falsification currently takes place with over-the-counter products rather than prescription medicine because inspections and regulations are less rigid. Although the majority of patients are aware that quality cannot be guaranteed, some products such as weight loss products, sexual products, or dietary supplements are sometimes purchased on the internet.

Second, an MOH official mentioned that ITS data could be used more effectively to prevent shortages and stock-outs. Although the current system is largely reactive, MOH aims to implement a proactive alarm system that provides a warning when the supply of a certain product goes below a specific threshold in a particular area. Such warnings will enable the system to procure medicine more rapidly and to prevent drug shortages more successfully.

**DISCUSSION**

Several countries and regions have attempted to introduce full pharmaceutical track and trace systems. Turkey was the first to succeed. This study aimed to elucidate the factors underpinning the success of this technological innovation. We find that the drivers of success were more political and economic than technological.

China, which has considerable experience and capability in implementing large technological programs in its health sector, suspended plans to introduce pharmaceutical track and trace in 2016 after facing considerable resistance from medicine manufacturers. Industry was concerned that the linear barcode system proposed instead of a DataMatrix code would create a large footprint on the medicine package to capture the required data. Further, there was concern that the requirement that all barcodes be printed only by the government would result in a burdensome and costly procedure for manufacturers. As discussions with stakeholders continue, the target data for implementation has been pushed back to 2022.

The United States provides another example in which difficulty in adequately aligning incentives for all key actors has led to slow adoption of full traceability. Industry has not fully complied with the phased implementation foreseen in the 2013 Drug Supply Chain Security Act. Implementation of the act is expected to take a full decade.4,34 How was Turkey, a middle-income country with no great tradition of technological innovation, able to succeed where others stumbled? The most critical element was the combination of a widely recognized problem and political determination to solve it.

The winners of the 2002 elections in Turkey sought to establish political legitimacy through programs that delivered benefits to a broad swath of citizens. One of these benefits was universal health coverage delivered through a single-payer state institution. When high levels of fraud threatened the sustainability of this coverage, politicians threw their weight behind an ambitious technological solution within an improbably tight timeframe. The state controlled access to a large and expanding pharmaceutical market, and manufacturers who wanted to sell into that market had to play ball. A generous benefit package greatly reduced out-of-pocket spending on medicines. Together with a prohibition on internet sales of prescription products, the benefit package removed any incentive for patients to purchase products from the unregulated supply chain. At the same time, the reimbursement system obliged pharmacists to bow to the will of the government; if they did not, they would not get paid. Together, these factors allowed for the widespread adoption of the system.

The successful implementation of the system was underpinned by another key factor: a willingness of the government, which was driving the process, to support flexible and adaptive solutions to problems identified during implementation. These work-arounds were not just technical; like all adaptive implementation, they also had a social component, encompassing human actions and relations.18 The Turkish state mainly focused on facilitating the social component, while other actors took responsibility for implementing the technical components of ITS.

Turkey’s centralized database allowed for verifying reimbursement data because its track and trace database was linked to the database of SGK, the single-payer state-owned reimbursement agency. This process enabled reducing fraud dramatically.9 Centralized databases rely heavily on the presence of sufficient technical capacity at the central level. If this capacity is lacking, outsourcing the development of the system to a software company, as in Turkey, might solve the problem, as long as security and privacy concerns of
stakeholders are addressed. In environments without political power emerging from a single-payer institution, the reimbursement landscape might be fragmented. The reimbursement agencies within a fragmented market, as well as the pharmaceutical industry, might oppose sharing and centralizing their competitive data more strongly.\textsuperscript{35,36} In these circumstances, setting up a distributed database that gives stakeholders more authority over their data might be more feasible. However, disadvantages of distributed databases include difficulty in governing and adapting the system because ownership of the data is not centralized.\textsuperscript{4,8}

To our knowledge, this study is the first to evaluate the emergence, implementation, and outcomes of ITS in Turkey, while focusing on the underlying political and economic factors. Since Turkey is the first country in the world to implement a full track and trace system, the implications of this study might be of particular interest to countries aiming to implement similar track and trace systems, including the EU member states, China, and the United States.

Limitations
The findings of this qualitative study could be strengthened through triangulation with quantitative data on the quality of medicine in the Turkish pharmaceutical market, the implementation costs of ITS, and the effect of ITS on public health. However, attempts to verify estimates provided by respondents with quantitative data from government or other formal sources proved unsuccessful. Future studies on the cost effectiveness of ITS might provide valuable insights.

For some categories of participants, we interviewed only a single key informant. However, triangulation of data provided by respondents from different sectors (e.g., manufacturers, wholesalers, technical agencies) in combination with further triangulation from literature increases our confidence in the validity and reliability of our study data.

RECOMMENDATIONS
The outcomes of our study show 3 main implications for countries aiming to implement pharmaceutical track and trace systems.

First, a track and trace system should be seen as a means to an end, rather than a goal in itself. To function, it must be underpinned by well-defined legislation and regulatory capacity, including laboratories and frequent GMP and GDP inspections. Without these, there is a risk of “garbage in equals garbage out.” In that case, track and trace may simply deliver a secure supply chain for poor-quality products.

Second, the incentives of all the stakeholders need to be aligned to successfully adopt the system. The role of the state should not be underestimated. It should both facilitate the implementation process with its political power, as well as provide sufficient leeway to adapt the system according to the needs of stakeholders. Countries lacking powerful political leadership might expect greater resistance to implementing pharmaceutical track and trace systems from stakeholders. This resistance is especially likely from stakeholders that bear the burden of upfront investment in technology and those that might benefit from gaps in the supply chain.

Third, countries/regions should aim to implement a full track and trace system. Although the benefits of track and trace systems are not universal and rely on the nature of the pharmaceutical system of the implementing country, point-of-dispense check systems, which exclude wholesalers and other middlemen, preclude some of the more important benefits of full track and trace. They do not provide data to flag regional shortages. Further, because they do not allow for traceability of products throughout the supply chain, such partial systems limit the ability to recall products. As a result, falsified products might circulate in a market for months without detection.\textsuperscript{3} This situation is especially true in the EU’s complex single-market supply chains. Although the European Medicines Verification System has added an antitampering device to the outer medicine package to prevent unlawful repackaging, nonreimbursed or over-the-counter products will remain vulnerable to falsification. Most importantly, in EU member states lacking closed supply chains, patients might buy products online that are less likely to be verified and may even be excluded from verification. Although accreditations, domain name verifications, and logos for online pharmacies exist, their effectiveness can still be undermined by a lack of consumer awareness, vulnerability to misuse, unavailability of certain types of products at accredited online pharmacies, and the attractiveness of cheaper options.\textsuperscript{37–39}

CONCLUSION
Although track and trace systems are sometimes presented as reproducible technical solutions to quality assurance in the supply chain, this study shows that the main drivers of success for ITS in
Turkey were highly dependent on the presence of a specific set of circumstances. These included political determination induced by reimbursement fraud, political power emerging from a single-payer institution that generated a substantial pharmaceutical market, reimbursement for verified dispensing and prescription, and flexibility to adapt the system according to the needs of stakeholders during implementation.

Despite ITS’s success in providing a clean regulated supply chain, it represents only part of the solution. ITS can only operate effectively if it is embedded in a pharmaceutical market where all legislative and regulatory components are in place.

Acknowledgments: The team is grateful to all interviewees participating in this study. We thank Adina-Loredana Nistor, Amalia Hasnida, Jiangying Xu, and Perneet Bourdillon-Estere for their constructive feedback and assistance during the analysis of the study data. We would also like to thank all professionals and researchers for their participation, expertise, and suggestions on the preliminary findings of this study during a meeting in London in April 2018.

Funding: This work was supported by the Wellcome Trust (209930_Z_17_Z) and Erasmus University (through the Research Excellence and Innovation grant).

Competing interests: None declared.

REFERENCES


Peer Reviewed

Received: February 21, 2020; Accepted: June 16, 2020; First published online: August 19, 2020


© Parmaksiz et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are properly cited. To view a copy of the license, visit http://creativecommons.org/licenses/by/4.0/. When linking to this article, please use the following permanent link: https://doi.org/10.9745/GHSP-D-20-00084
Measuring Service Quality and Assessing Its Relationship to Contraceptive Discontinuation: A Prospective Cohort Study in Pakistan and Uganda

Karen T. Chang,a Nirali M. Chakraborty,a Amanda M. Kalamar,b Waqas Hameed,c Ben Bellows,d Karen A Grépin,e Agha Xaher Gul,f Sarah E.K. Bradley,g Lynn M. Atuyambe,h Dominic Montagu,a

ABSTRACT

Background: The quality of contraceptive counseling that women receive from their provider can influence their future contraceptive continuation. We examined (1) whether the quality of contraceptive service provision could be measured in a consistent way by using existing tools from 2 large-scale social franchises, and (2) whether facility quality measures based on these tools were consistently associated with contraceptive discontinuation.

Methods: We linked existing, routinely collected facility audit data from social franchise clinics in Pakistan and Uganda with client data. Clients were women aged 15–49 who initiated a modern, reversible contraceptive method from a sampled clinic. Consented participants completed an exit interview and were contacted 3, 6, and 12 months later. We collapsed indicators into quality domains using theory-based categorization, created summative quality domain scores, and used Cox proportional hazards models to estimate the relationship between these quality domains and discontinuation while in need of contraception.

Results: The 12-month all-modern method discontinuation rate was 12.5% among the 813 enrolled women in Pakistan and 5.1% among the 1,185 women in Uganda. We did not observe similar associations between facility-level quality measures and discontinuation across these 2 settings. In Pakistan, an increase in the structural privacy domain was associated with a 60% lower risk of discontinuation, adjusting for age and baseline method (P < .001). In Uganda, an increase in the management support domain was associated with a 33% reduction in discontinuation risk, controlling for age and baseline method (P = .005).

Conclusions: We were not able to leverage existing, widely used quality measurement tools to create quality domains that were consistently associated with discontinuation in 2 study settings. Given the importance of contraceptive service quality and recent advances in indicator standardization in other areas, we recommend further effort to harmonize and simplify measurement tools to measure and improve contraceptive quality of care for all.

INTRODUCTION

Quality of care has long been considered an important factor influencing care-seeking behavior and health outcomes.1 The inclusion of universal health coverage (UHC) in the Sustainable Development Goals highlights the need for not only expanding access to...
care but also improving the quality of care.\textsuperscript{2} Aligned with these goals, the Family Planning 2020 (FP2020) partnership aims to reach 120 million additional modern contraceptive users by 2020\textsuperscript{3,4} by helping partner countries to engage new users and sustain voluntary contraceptive use among the large number of family planning adopters, particularly those at risk of discontinuing use in the first year.\textsuperscript{5} Women who wish to continue controlling their fertility but discontinue contraceptive use are considered “in need” and are more likely to have a mistimed or unwanted pregnancy.\textsuperscript{6} Implicit in this goal is the idea that intervening with improvements to quality of care can decrease high discontinuation rates and ultimately help women better realize their reproductive goals.\textsuperscript{3} Recent publications, including one by several authors of this paper, have documented a strong association between counseling quality, measured by either the Method Information Index (MII) or a variant and contraceptive continuation.\textsuperscript{7–9} Consensus on how best to define and measure facility-level quality of care in family planning has not yet been reached, however, and consensus is also lacking on the aspects of quality that may be important for reducing the risk of discontinuation while in need.

The framework established by Donabedian\textsuperscript{10} in 1988 outlines how quality of care can generally be defined and measured by linking structural components of settings where care is provided and processes of care provision to health outcomes. Within quality assessment of family planning clinical care, the framework developed by Bruce\textsuperscript{11} in 1990 identifies 6 components, including choice of contraceptive methods, information given to users, provider competence, client and provider relations, recontact and follow-up mechanisms, and an appropriate constellation of services. With the above frameworks as a foundation, quality assessment tools for family planning quality of care have proliferated in the past 2 decades. At times, prior studies aiming to assess the association between family planning quality of care and contraceptive discontinuation have focused on specific quality indicators, such as clients receiving their chosen method or high-quality counseling, to represent one component of the Bruce framework.\textsuperscript{12–14} Others have presented more comprehensive constructs of family planning quality, developing a large set of indicators meant to represent all components of these frameworks and relating these to contraceptive use or discontinuation.\textsuperscript{15–17} Well-known national surveys supported by the United States Agency for International Development (Demographic and Health Surveys [DHS], Service Provision Assessments [SPA]), World Health Organization (Service Availability Mapping, Service Availability and Readiness Assessment), and many others\textsuperscript{18} have all incorporated indicators to measure aspects of family planning quality. Studies have also tried to use indicators available in these surveys to assess the relationship between quality and contraceptive use or discontinuation.\textsuperscript{19–22} Disparate definitions of quality and different measurement tools make it difficult to compare findings across studies and to establish a consensus on the aspects of quality that are important for reducing the risk of discontinuation while in need.

In this study, we aimed to contribute to the literature on family planning quality of care in 2 ways. First, we examined whether data routinely collected from family planning facilities could be used to measure structural and process quality of care in a comparable way. We leveraged existing data in a practical, field-guided approach, hoping to align existing quality measurement tools with each other and with a facility-level quality framework (Figure 1). This framework, adapted from those of Donabedian\textsuperscript{10} and Bruce,\textsuperscript{11} was developed with input from a group of 19 experts comprising family planning providers, academics, donors, and research institutions convened at the Rockefeller Foundation conference facilities in Bellagio, Italy, in October 2015.\textsuperscript{24}

To achieve this first objective, we examined tools used by 2 of the largest global franchisors in family planning, Marie Stopes International (MSI) and Population Services International (PSI). Together, these organizations delivered over 10.8 million couple-years of protection in 2014.\textsuperscript{25} Both are considered social franchises: networks of private health care providers, connected through agreements to provide socially beneficial health services under a common franchise brand often with the goal of increasing the availability, affordability, and quality of services. In such arrangements, the franchisor typically provides training, commodities, and quality assurance, while the franchisees agree to provide franchised services, undergo audits, and adhere to price ceilings.\textsuperscript{26} Both MSI and PSI provide extensive training and technical support to family planning providers in their networks to ensure a standard set of high-quality services. Both franchises also use robust quality assessment tools to monitor the performance of private medical practitioners. We categorized items from each franchisor quality assessment audit tool into one of the 6 domains representing structural and process indicators, hoping that the disparate tools could
be aligned—and potentially streamlined—in a way that could reliably measure facility quality.

Second, we tested whether these harmonized facility indicators are related to service outcomes indicative of quality care by examining the relationship between the facility quality measures and contraceptive discontinuation while in need. For this objective, we linked the facility quality measures with data from clients of MSI and PSI clinics collected in a 12-month prospective cohort study of family planning clients in Pakistan and Uganda. The client study is described briefly in the current article and fully in Chakraborty et al.7 We used these linked data to assess whether the facility-based quality domains are associated with contraceptive discontinuation while in need among franchise clients in Pakistan and Uganda. Our goal in conducting this study was to find a streamlined set of comparable indicators linking facility quality of care to service outcomes. Such indicators would enable not just franchisees but also clinics with more limited resources, including those in the public sector, to monitor and improve family planning service quality.

**METHODS**

**Study Design**

Contraceptive method discontinuation data were collected longitudinally in Pakistan and Uganda using a prospective cohort design. Pakistan and Uganda were selected due to the high levels of unmet need for family planning and the presence of a strong research partner in each country. This study was conducted in collaboration with PSI’s Ugandan partner PACE-Uganda, which operates the ProFam franchise, and Marie Stopes Society (MSS) in Pakistan, which operates the Suraj franchise.26 By leveraging existing social franchises with high client volumes, we were able to quickly recruit a large number of women and to obtain data on quality of services delivered at health facilities in our sample.

**Facility Selection, Eligibility, Recruitment, and Follow-up**

A detailed description of facility selection, eligibility, recruitment, and follow-up can be found elsewhere.7 In brief, 75 Suraj social franchise centers and 30 ProFam clinics were sampled in Pakistan and Uganda, respectively. Women visiting these facilities were eligible to participate in this study if they had received a modern family planning method (male/female condom, pill, injectable, implant, intrauterine device, or emergency contraceptive) during the current visit and were (1) first-time users (defined as using contraception for the first time in their life), (2) current users switching to a different modern method, or (3) previous users with a 3-month or longer lapse who were returning to use. Women returning for resupply of a current method were not eligible for the study. Additionally, to be eligible in Uganda, women needed at least one mobile phone number at baseline for follow-up interviews.

Recruitment and baseline data collection took place in Pakistan from December 2016 to February 2017 and in Uganda from February to April 2017. In both countries, women leaving a study clinic were approached and screened for eligibility.
for the study. If eligible, they were invited to participate in the study, asked to provide written informed consent, complete an exit interview, and also provide information to allow them to be contacted 3, 6, and 12 months after the visit. In Uganda, the women were also consented for follow-up at 9 months.

Data Management

**Facility-Level Variables**

Health facility quality data were obtained from routine facility audit data sources from each social franchise organization. In Pakistan, continuous review processes including internal audits conducted by country programs as well as annual external assessments are completed using the MSS Quality Audit (QA) Checklist. This checklist includes a total of 305 indicators related to the full range of modern family planning services. Each item is scored as follows: 0, standard not in place; 1, standard partially achieved; or 2, standard fully achieved. The most recent QA data collected within 1 year of the start of enrollment for each clinic was used in this study. Variables in the MSS QA Checklist related to sterilization were excluded from the analysis because only women adopting reversible modern methods were eligible for inclusion. In this analysis, 215 facility quality variables were included. Additional non–family planning service delivery items in the checklist that were excluded in this analysis include those related to cervical cancer screening, postabortion care, and administrative support variables.

In Uganda, both annual internal audits and external audits at 2-year intervals are conducted using the PSI QA Family Planning Scorecard. This tool includes 5 QA standards and 21 indicators related to family planning services. Each item is scored as follows: 0, not acceptable; 1, competent; or 2, proficient. To ensure completeness of facility quality data for use in this study, QA assessments were conducted as part of study within 3 months of the start of enrollment. One of the 21 indicators from the PSI QA Family Planning Scorecard is evaluated at the national level and is not facility specific. Therefore, only 20 variables were included in this analysis. PSI also uses a number of checklists used to evaluate direct service provision, including infection prevention, client privacy, and continuity of care, that were not included in this analysis.

**Client-Level Variables**

A woman’s age was categorized as 15–24 years, 25–34 years, or 35+ years, and her primary baseline method was categorized as a short-acting versus a long-acting method. A household’s relative wealth was assessed using an asset index generated from the EquityTool, which makes use of a shortened list of country-specific assets benchmarked to the most recent DHS from each country (2012–2013 in Pakistan, 2016 in Uganda).

As in the MII-discontinuation analysis,^7^ the event of interest was discontinuation of a modern reversible method while in need, defined as a woman stopping contraceptive use without the intention to become pregnant while at risk of unwanted pregnancy.SPECIFICALLY, women were classified as discontinuing while in need of contraception if they self-reported stopping use of a modern method due to pregnancy (contraceptive failure), health concerns or side effects, desire for a more effective or convenient method, lack of continued access, financial reasons, or having a partner or family member who opposed further use. Women who reported one of these reasons for discontinuation at any of the follow-up interviews experienced the event and received no further follow-up. The definition of discontinuation while in need is related to, but distinct from, the concept of having an unmet need for contraception. Women who switched to another method were considered continuous modern method users. Those who reported stopping a previous method and starting a new modern method in the time between their previous and current visits were also considered continued users, regardless of the gap in time between stopping and starting their methods. Women who were lost to follow-up were right censored and assumed to be no different from those who remained in the study. Time to discontinuation was measured in days and treated as a continuous variable. In our analysis, although the maximum allowable time of follow-up was 360 days in Pakistan, time in Uganda was truncated to 300 days because no events took place in the final 60 days of the reporting period.

**Treatment of Missing Data**

**Facility-Level Variables**

In Pakistan, 38% of facility variables included values that were not observed and not applicable. To avoid dropping facilities from our analysis, observations that were not observed or not applicable were set to the mean of all other observations available for each variable. In Uganda, 10% of facility variables had missing values. Similarly, facility observations that were missing were assigned the mean of available observations for each variable.
Client-Level Variables

Complete baseline information was available for all women who were enrolled in the study. No dates of discontinuation were missing in Pakistan. In Uganda, for 12% of women who discontinued while in need, the date of discontinuation was missing or set to missing if the reported date fell outside 2 adjacent rounds of follow-up. Imputed dates were generated by randomly selecting a date between the 2 adjacent rounds of data collection for the woman. Of the 77 women in Uganda who reported method discontinuation, 5 were missing a reason for discontinuation, and we assumed they discontinued while in need. Of those who discontinued in Pakistan, all provided a reason for discontinuation.

Analytic Methods

Analysis of Facility Quality Variables

The aim of the analysis of facility quality variables was to categorize items from each franchisor quality tool into 1 of the 6 domains of interest, representing structural and process indicators of quality in the adapted framework (Figure 1). These domains include readiness for choice, readiness for management support, client-centered readiness, interpersonal skills, information provision, and technical competence. Readiness for choice measures the availability of contraceptive methods and tracer equipment. Readiness for management support includes having documentation of provider qualifications, training plans, and supervisory structure. Client-centered readiness refers to having the infrastructure that ensures privacy and confidentiality and the tools that help clients identify their preferred method. Interpersonal skills center on encouraging a positive relationship between clients and providers and ensuring clients are treated with dignity and respect. Information provision refers to the information exchanged between clients and providers that helps clients choose and use their contraception method, including information about advantages, disadvantages, and side effects of their chosen method as well as the alternative methods that are available. Technical competence involves adherence to quality standards, such as providers’ demonstrated competence in clinical techniques and ability to follow protocols and employ proper infection prevention techniques.

Analysis of facility-level variables was limited by the low facility to variable ratio, especially in Pakistan where we had only 75 facilities and 215 variables. This made collapsing variables into composite indicators a necessary first step in our analysis of facility quality data in Pakistan. A review of all 215 variables in a qualitative assessment of their importance to family planning quality led to the elimination of 27 variables. Examples of variables eliminated at this stage include those related to the exterior condition of the building and having a storage place for clients’ belongings. Additionally, 18 variables did not vary across the facilities and were excluded from the analysis. In creating composite indicators, conceptually related sets of variables were grouped together and tested to see if they reached a threshold of α > .65 for internal consistency. In this step, 154 individual variables were collapsed into 20 composite variables using facility quality data from the 75 facilities that participated in the study. In addition to these composite variables, 16 individual variables were retained in the Pakistan analysis, resulting in a total of 36 variables. In Uganda, 1 variable assessed the program only at the national level rather than facility-specific assessments and was dropped from the analysis. The remaining 20 individual variables were not collapsed into composite variables in Uganda.

Following the creation of composite indicators in Pakistan, the analytic steps taken for each country were similar. In both countries, items were assigned a priori to 1 of the 6 domains of our adapted quality of care framework. Factor analysis with orthogonal rotation was initially attempted in both study settings. However, items did not group together into domains that were meaningful or actionable. Therefore, a priori assignments of items to domains were used to group variables within a domain together based on an alpha greater than .65. If internal consistency of these a priori groupings fell below .65, the variable with the highest alpha consisting of all but the one item, suggesting poor fit with all other items, was removed from the grouping. In this case, the variable was then included separately in the model. Supplement 1 describes the types of variables included in the 6 quality domains in each country.

Analysis of Relationship Between Facility Quality and Contraceptive Discontinuation

To assess whether facility quality and discontinuation were correlated, we used survival analysis and Cox proportional hazard models with a shared frailty term to account for clustering by facility. The shared frailty term was significant in our analysis of data from Pakistan and not significant in Uganda. Therefore, in our analysis of Uganda...
data, we employed robust standard errors in the Cox proportional hazard models. Discontinuation rates were estimated from Kaplan-Meier survival curves. Both univariate models with each individual domain and a model with all 6 quality domains were run. In the full model with all 6 domains, multicollinearity between domains was assessed using variance inflation factors. Domains with a variance inflation factor greater than 4 were removed from the full domain model. We tested potential time-invariant covariates in each country including age, relative wealth group, parity, education, method type at baseline (short or long acting), and user type at baseline (first time user, returning to contraception after a lapse in use, method switcher). Regression models incorporated the Efron approximation for ties. We tested assumptions of proportionality for each covariate both graphically and numerically. Covariates that were significant at \( P < .10 \) in either the unadjusted univariate model, including a single quality domain, or the unadjusted full model, including all 6 domains, were considered for the final adjusted model. Models were assessed using a likelihood ratio test and comparing Akaike information criterion values. Results present the parsimonious Cox proportional hazards model for both contexts, adjusted for covariates that met significance criteria in at least one country. Additional sensitivity analyses were also completed to investigate the relationship between discontinuation and individual quality variables of interest, rather than domains, including contraceptive availability, structural privacy, confidentiality, and contraceptive counseling.

## RESULTS

The baseline demographic and reproductive health characteristics of the study participants have been summarized elsewhere.\(^7\) In brief, 813 and 1,185 women were enrolled in Pakistan and Uganda, respectively. Clients in Pakistan were older, had higher parity, were less educated, and belonged to lower relative wealth quintiles compared with clients in Uganda. As described by Chakraborty et al,\(^7\) important differences in the method adopted at baseline were observed between the 2 settings. In Pakistan, 43.1% of clients reported using the intrauterine device compared with 23.3% in Uganda. No clients took up implants in Pakistan, whereas 36.4% of clients took up an implant in Uganda. In Pakistan, 24.5% of clients sought an injectable compared with 28.3% in Uganda. For the pill, 18.3% of clients in Pakistan accepted it at baseline compared with 10.3% in Uganda. Additionally, 14.2% of clients reported using male condoms in Pakistan versus only 1.8% in Uganda. At the time of enrollment, the 75 facilities in Pakistan had been part of the MSS franchise for 3 years on average, had 1 provider on site on any given day, and had 22 clients per week. The most qualified provider on site at the time of enrollment was most frequently a lady health visitor (45%), followed by a community midwife (41%), and 76% of the facilities were owner operated. The 30 facilities in Uganda averaged 4 years in the franchise, had 1.7 providers on site on any given day, and had a client volume of 16 per week. During enrollment, the most qualified provider was most often an enrolled midwife (50%) or registered midwife (23%), and most facilities (80%) were owner operated.

Table 1 presents the summary statistics of the 6 domains of facility quality, according to country. In Pakistan, technical competence had the highest number of items, 11 (\( \mu=15.7, \sigma=4.0, \min=3.6, \max=21.5 \)), followed by readiness for choice, 7 (\( \mu=12.2, \sigma=1.9, \min=5.4, \max=14.0 \)). The readiness for management support domain comprised 6 items, followed by client-centered readiness and interpersonal skills (4 items each), and information provision (3 items). The structural privacy variable did not group well with variables in other domains at an alpha level of .65, and it was included in the Pakistan model separately. In Uganda, the structural privacy variable was included in the client-centered readiness domain. Of the 20 items from the quality assessment checklist included in this analysis, 7 items were grouped into technical competence (\( \mu=11.8, \sigma=2.4, \min=7.0, \max=14.0 \)) and 5 in readiness for management support (\( \mu=8.9, \sigma=1.8, \min=5.0, \max=10.0 \)). Three items were included in client-centered readiness, 4 items in information provision, and a single item for readiness for choice. The quality assessment in Uganda had no indicators related to interpersonal skills. Compared with Uganda, Pakistan had a higher number of items in each quality domain, except for the information provision domain, where 3 items were included in Pakistan versus 4 in Uganda. Across both countries, the mean and range of quality scores in each domain were proportional to the number of items it contained. We observed relatively lower variation in the scores in Uganda as opposed to Pakistan.

Categorization of variables was influenced by the number of available variables in the Pakistan data, allowing for the creation of more specific composite variables. For example, in Uganda, a single indicator asks whether clients have access...
to a range of contraceptive methods. In Pakistan, 4 variables plus 3 composite indicators are included in the same domain (Supplement 1). A composite variable that relates to processes in place for infection prevention consists of 4 input variables, and one related to structural aspects of infection prevention (cleanliness, waste disposal, hand washing facilities) comprises 12 input variables. Variables were not always categorized in the same domain across the 2 countries. In Uganda, the presence of infection prevention equipment and supplies was assessed with the presence of "required" medical equipment and was categorized under the client-centered readiness domain. Details on what constitutes required equipment is not specified in the indicator.

Figure 2 presents the overall cumulative probability of women discontinuing use of their modern method while in need in Pakistan and Uganda. Discontinuation rates differed between the 2 countries, with a higher cumulative probability of modern method discontinuation while in need at 360 days of follow-up in Pakistan (12.5%), which had a substantially greater proportion of short-acting method users, compared with Uganda (5.1%) at the end of 300 days.

The unadjusted and adjusted association between various domains of service quality and contraceptive discontinuation is shown by country in Tables 2 and 3. Adjusted models included only significant covariates. Model 1 assesses the association between each domain of quality and contraceptive discontinuation; Model 2 shows the hazard ratios of each domain while controlling for the effects of other domains of quality; and Model 3 presents the effect of domains that were significant in either Models 1 or 2 at the \( P = .10 \) level, adjusting for participants’ age and type of method (short or long acting) adopted at baseline.

In Pakistan (Table 2), 3 of 6 domains and the structural privacy variable demonstrated a significant effect on contraceptive discontinuation: interpersonal skills (hazard ratio [HR] \( \text{Model}_1 = 0.80; \) 95% confidence interval [CI] = 0.69, 0.92); technical competence (HR \( \text{Model}_1 = 0.92; \) 95% CI = 0.86, 0.98); information provision (HR \( \text{Model}_1 = 0.79; \) 95% CI = 0.68, 0.91); and structural privacy (HR \( \text{Model}_1 = 0.37; \) 95% CI = 0.24, 0.57) (Table 2, Model 1). Readiness for choice was marginally significant (HR \( \text{Model}_1 = 0.89; \) 95% CI = 0.77, 1.03). However, when all domains were included in the model simultaneously (Model 2), none of the quality domains exhibited any significant association with contraceptive discontinuation, except for the structural privacy variable, which remained strongly associated with discontinuation.

### TABLE 1. Descriptive Statistics and Cronbach’s Alpha Scores for Each Facility Quality Domain in Pakistan and Uganda

<table>
<thead>
<tr>
<th>Domain</th>
<th>Pakistan Range</th>
<th>No. of items (N=36)</th>
<th>Min</th>
<th>Max</th>
<th>Mean (µ)</th>
<th>SD (σ)</th>
<th>Alpha (α)</th>
<th>Uganda Range</th>
<th>No. of items (N=20)</th>
<th>Min</th>
<th>Max</th>
<th>Mean (µ)</th>
<th>SD (σ)</th>
<th>Alpha (α)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Readiness for choice</td>
<td>0–14</td>
<td>7</td>
<td>5.42</td>
<td>14.00</td>
<td>12.20</td>
<td>1.94</td>
<td>.70</td>
<td>0–2</td>
<td>1</td>
<td>1.00</td>
<td>2.00</td>
<td>1.50</td>
<td>0.51</td>
<td>n/a&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Readiness for management support</td>
<td>0–12</td>
<td>6</td>
<td>1.77</td>
<td>11.83</td>
<td>8.02</td>
<td>2.88</td>
<td>.84</td>
<td>0–10</td>
<td>5</td>
<td>5.00</td>
<td>10.00</td>
<td>8.90</td>
<td>1.81</td>
<td>.89</td>
</tr>
<tr>
<td>Client-centered readiness</td>
<td>0–7</td>
<td>4</td>
<td>0.00</td>
<td>6.92</td>
<td>4.06</td>
<td>1.57</td>
<td>.66</td>
<td>0–6</td>
<td>3</td>
<td>3.00</td>
<td>6.00</td>
<td>5.07</td>
<td>1.17</td>
<td>.79</td>
</tr>
<tr>
<td>Interpersonal skills</td>
<td>0–8</td>
<td>4</td>
<td>0.67</td>
<td>8.00</td>
<td>5.49</td>
<td>1.94</td>
<td>.90</td>
<td>n/a&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technical competence</td>
<td>0–22</td>
<td>11</td>
<td>3.64</td>
<td>21.53</td>
<td>15.69</td>
<td>4.01</td>
<td>.80</td>
<td>0–10</td>
<td>7</td>
<td>7.00</td>
<td>14.00</td>
<td>11.83</td>
<td>2.41</td>
<td>.83</td>
</tr>
<tr>
<td>Information provision</td>
<td>0–6</td>
<td>3</td>
<td>0.00</td>
<td>6.00</td>
<td>2.97</td>
<td>1.91</td>
<td>.97</td>
<td>0–8</td>
<td>4</td>
<td>1.00</td>
<td>8.00</td>
<td>6.20</td>
<td>2.02</td>
<td>.65</td>
</tr>
<tr>
<td>Structural privacy</td>
<td>0–2</td>
<td>1</td>
<td>0.00</td>
<td>2.00</td>
<td>1.45</td>
<td>0.61</td>
<td>n/a&lt;sup&gt;c&lt;/sup&gt;</td>
<td>n/a&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** n/a, not applicable; SD, standard deviation.

<sup>a</sup> The structural privacy variable in Pakistan and the readiness for choice domain in Uganda each contained only 1 item; alpha not calculated.

<sup>b</sup> No variables were grouped into the interpersonal skills domain in Uganda.

<sup>c</sup> The structural privacy variable did not group well with variables in other domains in Pakistan. In Uganda, the structural privacy variable was included in the client-centered readiness domain.
FIGURE 2. Cumulative Probability of Modern Method Discontinuation Among Women in Need in Pakistan and Uganda

TABLE 2. Unadjusted and Adjusted Hazard Ratios for Discontinuation of Modern Contraception in Pakistan, by Quality Domain

<table>
<thead>
<tr>
<th>Facility Quality Domain</th>
<th>Unadjusted Univariate Model 1</th>
<th>Unadjusted Full Model 2</th>
<th>Adjusted Reduced Model</th>
<th>Use short-acting method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Readiness for choice</td>
<td>0.89 (0.77, 1.03)</td>
<td>0.94 (0.81, 1.10)</td>
<td>0.92 (0.80, 1.05)</td>
<td>1.80 (1.14, 2.86)</td>
</tr>
<tr>
<td>Readiness for management support</td>
<td>1.02 (0.90, 1.15)</td>
<td>0.91 (0.76, 1.09)</td>
<td>0.93 (0.81, 1.09)</td>
<td>1.60 (1.02, 2.83)</td>
</tr>
<tr>
<td>Client-centered readiness</td>
<td>1.14 (0.90, 1.44)</td>
<td>1.23 (0.92, 1.65)</td>
<td>1.23 (0.92, 1.65)</td>
<td>1.70 (1.02, 2.83)</td>
</tr>
<tr>
<td>Interpersonal skills</td>
<td>0.80 (0.69, 0.92)</td>
<td>0.97 (0.76, 1.25)</td>
<td>0.97 (0.76, 1.25)</td>
<td>1.70 (1.02, 2.83)</td>
</tr>
<tr>
<td>Technical competence</td>
<td>0.92 (0.86, 0.98)</td>
<td>1.00 (0.90, 1.11)</td>
<td>1.00 (0.90, 1.11)</td>
<td>1.70 (1.02, 2.83)</td>
</tr>
<tr>
<td>Information provision</td>
<td>0.79 (0.68, 0.91)</td>
<td>1.06 (0.79, 1.42)</td>
<td>1.06 (0.79, 1.42)</td>
<td>1.70 (1.02, 2.83)</td>
</tr>
<tr>
<td>Structural privacy</td>
<td>0.37 (0.24, 0.57)</td>
<td>&lt;.001</td>
<td>0.36 (0.18, 0.71)</td>
<td>0.36 (0.18, 0.71)</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; HR, hazard ratio; ref, reference category; SE, standard error.

a Adjusted for participants’ age and short-acting versus long-acting method use at baseline.

b Theta (shared frailty term): t=0.88, SE=0.33, P<.001.
c Theta (shared frailty term): t=0.99, SE=0.35, P<.001.
d Theta (shared frailty term): t=0.98, SE=0.34, P<.001.
e Theta (shared frailty term): t=0.68, SE=0.28, P<.001.
f Theta (shared frailty term): t=0.72, SE=0.30, P<.001.
g Theta (shared frailty term): t=0.63, SE=0.29, P<.001.
h Theta (shared frailty term): t=0.47, SE=0.23, P<.001.
i Theta (shared frailty term): t=0.44, SE=0.22, P=0.01.
j Theta (shared frailty term): t=0.55, SE=0.27, P<.001.
These results remained consistent after controlling for participant characteristics (not shown). In the reduced Model 3, with controlling for participants’ characteristics, structural privacy remained the only significant predictor of contraceptive discontinuation: a one-point increase in the score was associated with a 60% lower risk of discontinuation.

In Uganda, none of the quality domains showed any significant association with contraceptive discontinuation in the univariate model (Table 3, Model 1). With controlling for other domains of quality (Model 2), the measure of readiness for management support became significant, and remained significant when controlling for participants’ age and type of method in Model 3: a one-unit increase in the score was associated with a 33% lower likelihood of discontinuation.

Interestingly, when other domains were accounted for (Model 2), technical competence showed a counterintuitively negative association with discontinuation: a one-unit improvement in the technical competence score was associated with a 47% higher risk of discontinuation. However, this relationship was only marginally significant after the inclusion of age and type of method (HRModel3= 1.32; 95% CI= 0.97, 1.78) (Model 3). It is worthwhile to note that 1 of the 7 variables in the technical competence domain is specific to long-acting reversible contraceptive services and provision, but the majority of discontinuation in Uganda was among injectable users, making this association difficult to interpret.

Sensitivity analyses of individual quality variables of interest did not produce consistent results across the 2 countries. Contraceptive availability, defined as method availability on site or through referral in Pakistan and having access to a range of methods or information about where to obtain a method in Uganda, was not significantly associated with discontinuation in univariate analysis or after adjusting for age and short-acting versus long-acting method in either setting (Supplement 2). In Pakistan, (1) having a setting that offered client privacy, (2) keeping client cases confidential, and (3) counseling on key points, benefits, and side effects of methods at baseline were each significantly protective of discontinuation (1: HRadj= 0.38; 95% CI= 0.25, 0.58), (2: HRadj= 0.47; 95% CI= 0.29, 0.76), and (3: HRadj= 0.49; 95% CI= 0.32, 0.74), respectively, controlling for age and method type. However, in Uganda, none of these variables had a significant relationship with discontinuation, adjusting for age and method type (1: HRadj= 0.89; 95% CI= 0.42, 1.88), (2: HRadj= 0.73; 95% CI= 0.34, 1.59), and (3: HRadj= 1.29; 95% CI= 0.66, 2.51).

**TABLE 3.** Unadjusted and Adjusted Hazard Ratios for Discontinuation of Modern Contraception in Uganda, by Quality Domain

<table>
<thead>
<tr>
<th>Facility Quality Domain</th>
<th>Unadjusted Univariate Model Model 1</th>
<th>Unadjusted Full Model Model 2</th>
<th>Adjusted Reduced Modela Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>P-value</td>
<td>HR (95% CI)</td>
</tr>
<tr>
<td>Readiness for choice</td>
<td>0.79 (0.42, 1.47)</td>
<td>.46</td>
<td>0.59 (0.28, 1.23)</td>
</tr>
<tr>
<td>Readiness for management support</td>
<td>0.96 (0.82, 1.13)</td>
<td>.64</td>
<td>0.60 (0.42, 0.86)</td>
</tr>
<tr>
<td>Client-centered readiness</td>
<td>1.05 (0.78, 1.41)</td>
<td>.76</td>
<td>1.30 (0.92, 1.82)</td>
</tr>
<tr>
<td>Interpersonal skills</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Technical competence</td>
<td>1.07 (0.91, 1.25)</td>
<td>.43</td>
<td>1.47 (1.07, 2.02)</td>
</tr>
<tr>
<td>Information provision</td>
<td>1.03 (0.89, 1.18)</td>
<td>.73</td>
<td>1.05 (0.87, 1.25)</td>
</tr>
<tr>
<td>Age category (35–49, ref)</td>
<td>—</td>
<td>—</td>
<td>1.93 (0.62, 6.03)</td>
</tr>
<tr>
<td>15–24</td>
<td>2.54 (1.01, 6.39)</td>
<td>.05</td>
<td>7.64 (3.91, 14.92)</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; HR, hazard ratio; ref, reference category.

a Adjusted for participants’ age and short-acting versus long-acting method use at baseline.

**DISCUSSION**

We approached this study with 2 objectives. First, to determine if facility quality could be measured in a consistent way using existing tools from...
2 large international social franchise organizations that provide millions of couple-years of protection annually. That is, could we align structural and process quality domains in a consistent way across the 215 variables routinely collected in the MSS Quality Audit Checklist and the 21 facility-level indicators collected with PSI’s QA family planning scorecard? Second, to understand which, if any, aspects of quality as measured by these organizations are related to all-method discontinuation while in need.

In this field-based approach, we chose to leverage 2 different existing quality measurement tools to engender more confidence in potential findings. If similar associations were observed using tools of 2 large franchises, we may be more confident that those domains of quality matter in terms of reducing clients’ risk of discontinuation while in need. Beyond addressing the resource-intensive requirements of existing quality measurement, demonstrating a simple, robust measure to link quality of care and access to services motivated this study. Our rationale was that if similarities between the MSI and PSI measurement systems could be identified, then quality measurement may be done more efficiently by requiring only a key set of important quality indicators for these and other organizations with fewer resources. Furthermore, understanding if quality is related to a service outcome, such as discontinuation, may help draw attention and resources to quality measurement in the public sector, where resources are scarce, while achieving shared goals of improved quality and reduced unmet need. We recognize that receiving good quality of care is a human right, and we aim to provide data that are actionable to better guide program managers in choosing which indicators to focus on.

Our study did not find similar patterns in facility-level quality measures and discontinuation between these 2 settings. In Pakistan, women who received methods from facilities with higher structural privacy scores were 60% less likely to discontinue their methods, perhaps indicating that having privacy to discuss any questions or concerns about a method led to fewer concerns or more satisfaction with it and thus lower discontinuation. Analysis of individual quality variables demonstrated that structural privacy had a protective, but not statistically significant effect in Uganda. Readiness for management support was significantly associated with lower discontinuation in Uganda, but not in Pakistan. And in Uganda, technical competence was associated with higher levels of discontinuation when other quality measures were included in the model. This finding may be related to interactions between the items used to measure quality, which we were unable to assess due to sample size limitations. It may also be explained by the difference in training between the providers in each country. Although the most common provider cadre in Pakistan is the lady health visitor, who requires a minimum of only 8 years of schooling to be eligible for training, franchised providers in Uganda are secondary school or university graduates with further qualifications. More highly trained professionals may be less likely to account for client concerns. We also conducted sensitivity analyses to test for associations between individual quality items and discontinuation in case the domains were masking an effect; however, we did not observe consistent relationships across the 2 settings. Overall, we are unable to identify specific aspects of structural or process facility quality between the PSI and MSI tools that are consistently associated with continued contraceptive use, although the finding on structural privacy is encouraging and worthy of further investigation.

Comparisons between our findings and prior research investigating family planning quality and discontinuation are difficult because of the wide variety of measurement approaches and definitions of family planning service quality. Studies focusing on specific aspects of quality, such as client choice of method or improved counseling, reported significant positive relationships with sustained contraceptive use. A more comprehensive definition of quality was used in studies conducted in Senegal and the Philippines, where authors presented a single variable combined from a set of dichotomous items reflecting 5 different quality aspects, including whether a client’s needs were assessed and whether the client was provided information, was offered a choice in method, was treated well by her provider, and was linked to follow-up services. Both studies reported significant positive associations between the combined quality measures and continued contraceptive use; in the Philippines, having received information on all items versus none and having good versus poor interpersonal relations were each significantly associated with sustained modern contraceptive use. The authors explained that the selection of these variables was informed by Bruce’s conceptual framework. Two studies from the same project in Kenya identified elements of quality using factor analysis and investigated the relationships between each element and current modern contraceptive
Strengths of the current study were its prospective, clinic-based design and the ability to temporally establish the status of service quality prior to clients’ decision to adopt a contraceptive method. However, these strengths lead to difficulties in making comparisons with other studies from population-based surveys. For example, quality of care is not the focus of large national surveys, such as the DHS and SPA, and studies that utilize these datasets must limit their definitions of quality to the set of variables collected in the surveys. Additional quality variables may be of interest, but they are not collected in such surveys. Such surveys are retrospective, relying on self-reported contraceptive use months or years later that is subject to recall bias. Furthermore, linking household survey participants to facilities typically requires making assumptions about the facilities that participants attended within a specific catchment area. In reality, participants may have accessed other facilities of an entirely different level of service quality, potentially leading to tenuous conclusions about facility quality aspects that correlate with better outcomes. Retrospective reporting of contraceptive use in surveys also lacks the temporal requirement of assessing quality prior to a participant’s decision to adopt a contraceptive method.

Other strengths of this study include the follow-up of clients over 12 months, capturing the period during which the highest rates of discontinuation are expected to occur. Well-organized social franchise networks and existing infrastructure allowed for periodic follow-up visits that minimized overall attrition rates in both settings. Finally, the study leveraged existing quality assessment tools routinely used by social franchises. This approach provided a practical, and not simply theory-driven, opportunity to investigate whether tools currently being applied in programmatic settings could be simplified and streamlined for use in lower-resource facilities. This study does not provide evidence for how the tools can be streamlined, but it does suggest a need for greater comparability of measurement between service providers.

We faced several challenges in our attempts to align quality measures across settings with very different user profiles, quality measurement systems, and numbers of quality-related indicators. It is perhaps not surprising that the relationship between the 215 indicators from 75 facilities in Pakistan and discontinuation among older, higher-parity women primarily using short-acting methods is different from the relationship between 21 indicators from 30 facilities in Uganda and discontinuation among younger, lower-parity, more highly educated women from higher wealth quintiles who are primarily using long-acting methods. Beyond the differences in user profiles and the number of indicators, the systems used for measuring quality between social franchising organizations are quite different. Both organizations have devoted significant financial and human resources to establishing quality assurance systems and protocols, reflective of their desire to ensure clients receive high-quality care.

Given the resource intensiveness of the current tools used by organizations and the distinct indicators of these tools, standard quality measurement approaches could improve quality. Franchises have agreed to standard indicators of measurement in other areas (equity, impact, additional users). For quality, however, despite significant internal and donor-funded investments, organizations are not able to compare results, and the process of prioritizing indicators for management is unclear. In areas where a common measurement approach has been advanced, as has been done for equity, halo effects outside of the franchise sector have been seen. We hope this study contributes to ongoing discussions about comparable quality measures that eventually have an impact not only on franchises but also on other private and public sector actors.

Limitations
Our study is limited by the lack of direct comparability between the quality measurement tools in these 2 study settings. Although intentional, the use of program-specific tools meant that we needed to pragmatically apply analytic techniques to examine the differing quality measures. As noted, we did not have a sufficient sample size to use factor analysis, and thus relied on a theoretical approach to grouping variables into domains of quality. Franchised facilities, within a given setting, are
also relatively homogenous, by design. The PSI and MSS franchises have standards to adhere to, limiting variability in our measurement of quality. Finally, we assumed that quality assessments conducted within the year prior to the start of the study remained unchanged over the period of enrollment; however, it may be that quality had improved over time. Despite these limitations, however, we believe the findings are interesting, and merit follow-up.

**CONCLUSION**

Understanding the role of quality of care in increasing service use and improving health outcomes has been an area of renewed research focus in recent years, driven in part by the recognition that quality needs to be a key component of UHC. Many assessments have used the Donabedian and Bruce frameworks as starting points for understanding family planning quality. However, translating these frameworks into defined indicators for assessing quality in family planning programs has been a challenge. This challenge has underscored a need for standardized and simplified family planning quality measurement.

This study did not produce consistent findings across the 2 social franchise settings. We suspect one reason for this lack of consistency is the overall high level of quality, which made variations difficult to identify and correlations with better or worse quality difficult to discern. Despite these challenges, the findings from each site provide insights into whether existing tools used in the field can lead us to a common set of measures correlated with an outcome that is agreed to be modulated by quality. We conclude that such common measures can be found and that discontinuation while in need is the outcome researchers should assess measures against. However, given the heterogeneity in the current quality measurement approaches, further work is needed to harmonize measurement tools and performance indicators used by service delivery and research organizations. Although large donor-supported organizations can measure a wide range of aspects of care to ensure quality provision, the same is not true in most governmental programs, and in neither context are measures often assessed against changes in outcomes. To serve the needs of resource-constrained providers, research in a heterogeneous array of facilities may advance the field toward the simple standard measures needed for programmatic purposes. Ultimately, effective and pragmatic measurement of quality can make significant contributions to population health in the movement toward UHC.

**Acknowledgments:** The authors gratefully acknowledge the efforts of Marie Stopes Society of Pakistan and PACE Uganda in facilitating access to their franchised facilities for this study, as well as to the women who participated in the study.

**Competing interests:** None declared.

**REFERENCES**


Peer Reviewed

Received: March 12, 2020; Accepted: June 17, 2020; First published online: August 17, 2020


© Chang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are properly cited. To view a copy of the license, visit http://creativecommons.org/licenses/by/4.0/. When linking to this article, please use the following permanent link: https://doi.org/10.9745/GHSP-D-20-00105
Factors Associated With Delayed Contraceptive Implant Removal in Ethiopia

Elizabeth Costenbader, a Alice F. Cartwright, b Misti McDowell, c Berhane Assefa, d Meza Yirga Tejeji, c Eskindir Tenaw c

Key Findings
- Women receiving implant insertion at the community level were significantly more likely to report keeping their implant for more than 3 years.
- Task sharing only implant insertion and not removal services to community-based health worker cadres may have inadvertently led to extended implant use, particularly in rural areas or areas far from facilities.

Key Implications
- Governments and program planners should prioritize efforts to ensure that implant removal services are available at the same health system level as insertions.
- National stakeholders and implementing partners should consider including implant removal as part of the training for all health care providers who are trained in insertion.

ABSTRACT
Background: In 2009, the Government of Ethiopia initiated the implant scale-up initiative, which expanded contraceptive access by training health extension workers (HEWs) to insert single-rod etonogestrel contraceptive implants (Implanon) at rural health posts. Removals were provided by referrals to higher levels of the health system. However, little was known about whether women were getting their implants removed at the recommended 3-year postinsertion date or what barriers they faced to removal.

Methods: Between June and July 2016, 1,860 Ethiopian women, who had a 1-rod etonogestrel implant inserted by either an HEW or another health care provider between 3 and 6 years prior, were surveyed. We describe the characteristics of the sample and use multivariable logistic regression to predict factors associated with keeping implants inserted beyond 3 years.

Results: Women who had received their implants from HEWs were significantly more likely to report keeping them inserted for more than 3 years (adjusted odds ratio=2.50; 95% confidence interval=1.19, 5.24), compared with those who got their implant from another health care provider. Women who reported distance to the facility or transportation as a barrier were also significantly more likely to keep their implant for more than 3 years. Married and educated women were less likely to keep their implants for an extended duration. Among women who had their implant for 3 years or less, women who had had it inserted by an HEW were significantly more likely to report that the provider was unable or refused to provide removal as a barrier.

Discussion: Efforts to expand lower level and community-based access to contraceptive implants that do not ensure reliable access to removals at the same level as insertions may lead to women using implants beyond the recommended duration.

INTRODUCTION
Although Ethiopia has made considerable progress in increasing access to modern contraception over the past few decades, 1 Ethiopian women still have a high level of unmet need for contraception (20.6% for married women and 13.9% for all women), especially in rural areas (22.6% for rural married women compared with 13.1% in urban areas). 2 Part of the high unmet need among women in rural areas of Ethiopia may be driven by lack of access to facility-based health care in these areas. In 2009, to increase the national contraceptive prevalence and reduce unmet need for contraception, the Ethiopian Federal Ministry of Health...
Factors Associated With Delayed Contraceptive Implant Removal in Ethiopia

We wanted to assess whether having a 1-rod implant inserted by an HEW would be associated with keeping the implant longer than 3 years and what factors were associated with keeping it longer.

The Federal Ministry of Health (FMOH) launched an implant scale-up initiative to expand the family planning method mix and increase voluntary access to long-acting reversible contraceptives (LARCs). To achieve this goal, the FMOH made Implanon, a subdermal single-rod progestin-only implant, available at rural health posts by training health extension workers (HEWs) in its provision. Although clinical studies have found that 1-rod etonogestrel contraceptive implants may be effective for up to 5 years, since coming onto the market, Implanon (and its current analogue, Implanon NXT or Nexplanon) have only been approved for 3 years of effective use for pregnancy prevention before needing removal.

HEWs are government employees who are trained to provide a variety of services to rural communities. As part of the national Health Extension Program, which was launched in 2003 to increase access to preventive and curative health care services, HEWs provide services either at rural health posts or in residents’ homes. These preventive and curative services fall into 4 broad areas: health education and communication, hygiene and environmental sanitation, disease prevention and control, and family health services, which include family planning and adolescent reproductive health packages. HEWs counsel on the full range of methods and provide pills, condoms, and injectable contraceptives.

Before 2009, HEWs referred women who wished to obtain LARCs to a higher level of the health system. As part of the implant scale-up initiative, the FMOH collaborated with partner organizations to train HEWs in implant insertion. Notably, only the 1-rod contraceptive implant insertion was approved for task sharing to the HEW cadre. Substantial effort was also dedicated toward demand creation and ensuring the logistical commodity supply to expand accessibility. Currently, Implanon is the most widely used implant in Ethiopia and the most commonly used LARC method.

In the first 6 years of the scale-up in Ethiopia (2009–2015), more than 1.2 million single-rod etonogestrel contraceptive implant insertions occurred. During the same time period, only 37,175 Implanon removals were documented. HEWs were trained to counsel women on where to access removal services if they wanted the 1-rod contraceptive implant removed. Because implant removal involved making an incision, it was determined to be outside of the HEWs’ scope of practice, thus they were not trained in implant removal. Instead, as part of the implant scale-up initiative, additional health care providers at higher-level facilities were trained in removals, and mobile teams of these providers were sent to health posts and the community to provide removals on a periodic basis. Nonetheless, given the large number of implants provided by HEWs at the community level since the inception of the implant scale-up and the relatively few documented removals, information is limited regarding whether single-rod contraceptive implant users were able to get their implants removed at the recommended 3-year postinsertion date (or at any other time that they may have desired). Further, the barriers to removal that they may have faced are unknown.

For this study, we used data collected in 2016 to inform future planning of family planning service provision in Ethiopia and to assess whether women who received 1-rod contraceptive implants since the inception of the implant scale-up initiative had experienced any barriers to removal. The original parent study was further designed to inform recommendations to improve contraceptive implant service delivery. The main objectives of the current analysis were to test the hypothesis that single-rod contraceptive implant insertion by an HEW would be associated with keeping the implant for longer than 3 years and to assess which factors were most strongly associated with keeping implants inserted for longer than 3 years.

METHODS

Study Design and Sample
The data used for this analysis are from a survey conducted in the 4 regions in which the implant scale-up initiative with HEWs initially took place: Amhara; Oromia; Tigray; and Southern Nations, Nationalities, and Peoples’ Region (SNNPR). Data collection for the survey took place between June and July 2016 at 160 health posts and 40 health centers across those 4 regions.

Sampling Design
A stratified 3-stage cluster sample design was used to select survey respondents: woredas (districts), health facilities, and implant users were chosen as the first, second, and third stages of sampling, respectively. Before their selection, the woredas were stratified by urban and rural classifications. Subsequently, based on sample size power calculations, we determined that 12 woredas per region (for a total of 48 woredas) and 4 or 5 health facilities per woreda would be needed to reach our target of 200 health facilities and 2,000 women distributed equally across the 4 regions and proportionately across urban and rural strata.
We obtained lists of woredas from the Central Statistical Agency of Ethiopia within each of our 4 study regions, along with the estimated number of women of reproductive age. From each of these woreda health offices, we obtained lists of all their health facilities along with lists of the estimated number of 1-rod contraceptive implant users of reproductive age that were associated with each health facility. From these lists, we used a probability proportional to size sample selection scheme, taking the number of women of reproductive age within a woreda as a measure of size, to select woredas. The same sample selection scheme was used in selecting health facilities within woredas, using the number of single-rod contraceptive implant users as the measure of size. Finally, from the sampled health facilities, we obtained anonymized lists of implant users (who currently have or had a 1-rod contraceptive implant inserted between 3 and 6 years before the date of the interview). From these lists, contraceptive implant users were then chosen using equal probability systematic sample selection procedures.

Survey Procedures
The women eligible to be sampled for this study were identified in each health facility (health post and health center) from the family planning registers and family folders. These records were only accessed by clinic employees. The clinic employees first generated a list of all women who met the study eligibility criteria: willing and able to give informed consent for participation, age between 18 and 49, living in the catchment area of the health facility, and had a 1-rod etonogestrel contraceptive implant inserted between 3 and 6 years before the date of interview, along with an associated client identification number. An anonymized list of client identification numbers was shared with the study data collectors to sample a subset of women to be interviewed.

After a woman was selected, a health provider contacted each woman and requested that she come to the health post or health center on a specific day for the interview. Because the majority of women did not have phones, HEWs contacted the selected women in person either at the health facility or at the woman’s home. Once in a private location, the HEW described the study in general terms and asked if the woman was interested in learning more about the study. For those who were interested, HEWs provided additional details about the study to confirm eligibility and gave the women a date and time to come to the health post for study enrollment and to take part in an interview. Upon arrival at the health post, members of the study team described the study, reconfirmed eligibility, obtained informed consent, and administered the survey.

The survey tool included approximately 36 questions pertaining to participant demographics, participant contraceptive implant use (including when and by whom the implant was inserted, what information was provided by the provider on contraceptive implant use and removal, whether the implant was removed and if so, when, by whom, and the reason for removal, and any barriers/challenges to removal), and any subsequent/current family planning use. If a woman had had more than 1 implant inserted during the 3- to 6-year time frame, the interviewer was instructed to ask her about the latest implant. The interviews took place in a private room at the health facility, and if no private room was available, a place outside was used where they could not be overheard. The survey was paper based; conducted in either Amharic, Oromifa, or Tigrigna language, as applicable; and lasted approximately 20 minutes. Participants were reimbursed 50 birr (US$2.40) for their travel to and from the health facility for the interview. This study was reviewed and approved by the institutional review boards at the Ethiopian Public Health Institute and FHI 360 (Protection of Human Subjects Committee) in May 2016.

Analysis
The survey data were entered into EpiData 3.1 and then exported to SAS for analysis. To reduce biases introduced by the lack of reliable sampling frames and interviews refused or lost by primary sampling unit and strata, we constructed sampling weights to apply to the survey data. Key outcome indicators in the data were reviewed using bivariate analyses, including chi-square tests for significance. We then tested the bivariate relationship between keeping an implant beyond the recommended timeframe (3 years after insertion) and selected independent variables that we expected to affect implant removal in Ethiopia using design-adjusted chi-square tests. Based on the efforts to train HEWs to provide 1-rod contraceptive implant insertions but not removals, we hypothesized a priori that implant insertion by an HEW would be associated with keeping the implant for longer than 3 years. Informed by the results of the bivariate analyses, we subsequently developed a multivariable logistic model, including interactions between

Based on the efforts to train HEWs to provide implant insertions but not removals, we hypothesized a priori that implant insertion by an HEW would be associated with keeping the implant for longer than 3 years.
sociodemographic characteristics and the type of provider that inserted the implant. We ultimately selected a model based on model fit and collinearity assessments.

**RESULTS**

A total of 1,860 Ethiopian single-rod contraceptive implant users completed the survey (Figure 1). Notably, due to the security issues in Amhara during data collection, only 10 woredas were sampled instead of the planned number of 12. Table 1 shows the sociodemographic characteristics of contraceptive implant users by the provider that originally inserted their implant. Overall, women who had their implant inserted by an HEW were slightly older, significantly more likely to report having no education, to be married, and to live in Amhara or SNNPR compared with women surveyed who had had their implant inserted by another health provider at a health center. Those women whose 1-rod contraceptive implant had originally been inserted by an HEW were also significantly more likely to report no current contraceptive use and significantly less likely to report current use of a contraceptive implant. There was no significant difference in reasons provided for getting the contraceptive implant inserted (e.g., for spacing or limiting pregnancies) between those who had received the implant from an HEW versus another health provider (results not shown).

Although women who had their single-rod contraceptive implant inserted by an HEW were equally likely as women who had their implant inserted by another health care provider to report being informed about possible side effects of the implant and when it should be removed, they were slightly, but significantly more likely to report that they could not remember if they were told where to get their implant removed (Table 2). However, the overwhelming majority of women in both groups reported having been told when their implant should be removed (95% and 96%, respectively) as well as advised where they could get it removed (93% and 93%, respectively), and just over three-quarters (76% and 77%, respectively) reported being informed about possible side effects of a 1-rod contraceptive implant. In terms of duration of use of the contraceptive implant, over one-quarter (26%) of women who had had their implant inserted by an HEW reported using it for more than 3 years (21%) or that it was still inserted (3–6 years after insertion) (5%), compared

![Figure 1. Number of Implanon Users Surveyed in Ethiopia, by Region](image)
TABLE 1. Demographic Characteristics of Study Participants by Provider That Inserted Implanon<sup>a</sup>

<table>
<thead>
<tr>
<th></th>
<th>HEW (n=1,346)</th>
<th>Other Health Provider (n=514)</th>
<th>Total (N=1,860)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age&lt;sup&gt;b&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–24</td>
<td>166 (12.3)</td>
<td>86 (18.0)</td>
<td>252 (14.4)</td>
<td>.0017</td>
</tr>
<tr>
<td>25–34</td>
<td>704 (57.3)</td>
<td>283 (50.7)</td>
<td>987 (54.8)</td>
<td></td>
</tr>
<tr>
<td>35–44</td>
<td>397 (27.2)</td>
<td>133 (30.5)</td>
<td>530 (28.5)</td>
<td></td>
</tr>
<tr>
<td>45–49</td>
<td>76 (3.2)</td>
<td>12 (0.8)</td>
<td>88 (2.3)</td>
<td></td>
</tr>
<tr>
<td><strong>Level of education</strong></td>
<td></td>
<td></td>
<td></td>
<td>.0012</td>
</tr>
<tr>
<td>No education</td>
<td>807 (60.3)</td>
<td>214 (46.0)</td>
<td>1,021 (55.0)</td>
<td></td>
</tr>
<tr>
<td>Read and write</td>
<td>82 (6.2)</td>
<td>25 (4.7)</td>
<td>107 (5.6)</td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>374 (28.2)</td>
<td>187 (38.9)</td>
<td>561 (32.2)</td>
<td></td>
</tr>
<tr>
<td>Secondary</td>
<td>80 (5.0)</td>
<td>71 (8.6)</td>
<td>151 (6.4)</td>
<td></td>
</tr>
<tr>
<td>Tertiary</td>
<td>3 (0.3)</td>
<td>17 (1.8)</td>
<td>20 (0.8)</td>
<td></td>
</tr>
<tr>
<td><strong>Religion</strong></td>
<td></td>
<td></td>
<td></td>
<td>.84</td>
</tr>
<tr>
<td>Orthodox</td>
<td>712 (42.0)</td>
<td>244 (36.2)</td>
<td>956 (39.8)</td>
<td></td>
</tr>
<tr>
<td>Protestant</td>
<td>328 (37.7)</td>
<td>123 (36.8)</td>
<td>451 (37.4)</td>
<td></td>
</tr>
<tr>
<td>Muslim</td>
<td>276 (18.5)</td>
<td>136 (23.7)</td>
<td>412 (20.5)</td>
<td></td>
</tr>
<tr>
<td>Catholic/other</td>
<td>30 (1.8)</td>
<td>11 (3.3)</td>
<td>41 (2.4)</td>
<td></td>
</tr>
<tr>
<td><strong>Marital status</strong></td>
<td></td>
<td></td>
<td></td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Married</td>
<td>1,219 (93.1)</td>
<td>443 (85.7)</td>
<td>1,662 (90.3)</td>
<td></td>
</tr>
<tr>
<td>Divorced/widowed/separated</td>
<td>110 (5.5)</td>
<td>59 (10.3)</td>
<td>169 (7.3)</td>
<td></td>
</tr>
<tr>
<td>Never married</td>
<td>17 (1.4)</td>
<td>12 (4.0)</td>
<td>29 (2.4)</td>
<td></td>
</tr>
<tr>
<td><strong>Employment status</strong>&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td>.42</td>
</tr>
<tr>
<td>Farm work</td>
<td>844 (54.7)</td>
<td>225 (67.3)</td>
<td>1,069 (59.4)</td>
<td></td>
</tr>
<tr>
<td>Housewife</td>
<td>266 (27.1)</td>
<td>109 (14.0)</td>
<td>375 (22.2)</td>
<td></td>
</tr>
<tr>
<td>Merchant</td>
<td>162 (10.9)</td>
<td>100 (10.9)</td>
<td>262 (10.9)</td>
<td></td>
</tr>
<tr>
<td>Public servant/other/student/not employed</td>
<td>74 (7.3)</td>
<td>80 (7.9)</td>
<td>154 (7.5)</td>
<td></td>
</tr>
<tr>
<td><strong>Region</strong></td>
<td></td>
<td></td>
<td></td>
<td>.0059</td>
</tr>
<tr>
<td>Tigray</td>
<td>379 (10.5)</td>
<td>101 (12.7)</td>
<td>480 (11.3)</td>
<td></td>
</tr>
<tr>
<td>Oromia</td>
<td>274 (24.8)</td>
<td>235 (63.7)</td>
<td>509 (39.4)</td>
<td></td>
</tr>
<tr>
<td>Amhara</td>
<td>304 (18.4)</td>
<td>96 (5.5)</td>
<td>400 (13.5)</td>
<td></td>
</tr>
<tr>
<td>SNNP&lt;sup&gt;r&lt;/sup&gt;</td>
<td>389 (46.4)</td>
<td>82 (18.1)</td>
<td>471 (35.8)</td>
<td></td>
</tr>
<tr>
<td><strong>Contraceptive method(s) currently using</strong>&lt;sup&gt;e&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>511 (39.7)</td>
<td>152 (20.1)</td>
<td>663 (32.3)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Implanon</td>
<td>332 (25.0)</td>
<td>176 (49.1)</td>
<td>508 (34.0)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Other implant</td>
<td>15 (2.0)</td>
<td>7 (1.8)</td>
<td>22 (1.9)</td>
<td>.94</td>
</tr>
<tr>
<td>Injectable</td>
<td>432 (28.8)</td>
<td>145 (24.3)</td>
<td>577 (27.1)</td>
<td>.53</td>
</tr>
</tbody>
</table>

<sup>a</sup>TABLE continued...

Continued
with just over 10% (11% more than 3 years and 1% still inserted) of those who had had implant insertion performed by another provider, a statistically significant difference.

Five percent of survey respondents reported still having their contraceptive implant at the time of the survey, although it was past the recommended removal date (Table 2 footnote). The

**TABLE 1.** Continued

<table>
<thead>
<tr>
<th>Method of Family Planning</th>
<th>HEW (n=1,346)</th>
<th>Other Health Provider (n=514)</th>
<th>Total (N=1,860)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pills</td>
<td>24 (1.9)</td>
<td>23 (3.4)</td>
<td>47 (2.5)</td>
<td>.35</td>
</tr>
<tr>
<td>IUD</td>
<td>26 (2.1)</td>
<td>9 (1.1)</td>
<td>35 (1.8)</td>
<td>.36</td>
</tr>
<tr>
<td>Condom/calendar/other</td>
<td>7 (0.6)</td>
<td>4 (0.2)</td>
<td>11 (0.4)</td>
<td>.13</td>
</tr>
</tbody>
</table>

Abbreviations: HEW, health extension worker; IUD, intrauterine device; SNNPR, Southern Nations, Nationalities, and Peoples’ Region.

* Frequencies and weighted percentages reported.

**TABLE 2.** Characteristics of Implanon Insertion Experience and Use by Provider That Inserted Implanon

<table>
<thead>
<tr>
<th></th>
<th>HEW (n=1,346)</th>
<th>Other Health Provider (n=514)</th>
<th>Total (N=1,860)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Informed about possible side effects of Implanon</td>
<td></td>
<td></td>
<td></td>
<td>.61</td>
</tr>
<tr>
<td>Yes</td>
<td>1,026 (76.8)</td>
<td>404 (77.3)</td>
<td>1,430 (77.0)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>300 (21.9)</td>
<td>105 (22.2)</td>
<td>405 (22.0)</td>
<td></td>
</tr>
<tr>
<td>Can’t remember/no response</td>
<td>20 (1.3)</td>
<td>5 (0.5)</td>
<td>25 (1.0)</td>
<td></td>
</tr>
<tr>
<td>Told when Implanon should be removed</td>
<td></td>
<td></td>
<td></td>
<td>.46</td>
</tr>
<tr>
<td>Yes</td>
<td>1,276 (95.0)</td>
<td>490 (96.5)</td>
<td>1,766 (95.6)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>61 (4.1)</td>
<td>21 (3.2)</td>
<td>82 (3.7)</td>
<td></td>
</tr>
<tr>
<td>Can’t remember/no response</td>
<td>9 (0.9)</td>
<td>3 (0.3)</td>
<td>12 (0.7)</td>
<td></td>
</tr>
<tr>
<td>Advised where they could get Implanon removed</td>
<td></td>
<td></td>
<td></td>
<td>.021</td>
</tr>
<tr>
<td>Yes</td>
<td>1,230 (93.0)</td>
<td>482 (93.1)</td>
<td>1,712 (93.1)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>105 (6.2)</td>
<td>31 (6.9)</td>
<td>136 (6.5)</td>
<td></td>
</tr>
<tr>
<td>Can’t remember/no response</td>
<td>11 (0.7)</td>
<td>1 (0.0)</td>
<td>12 (0.4)</td>
<td></td>
</tr>
<tr>
<td>Duration of Implanon use</td>
<td></td>
<td></td>
<td></td>
<td>.0010</td>
</tr>
<tr>
<td>&lt;1 year</td>
<td>30 (1.7)</td>
<td>19 (2.0)</td>
<td>49 (1.8)</td>
<td></td>
</tr>
<tr>
<td>1 year to &lt;3 years</td>
<td>251 (16.7)</td>
<td>100 (14.7)</td>
<td>351 (16.0)</td>
<td></td>
</tr>
<tr>
<td>3 years</td>
<td>781 (55.3)</td>
<td>311 (70.9)</td>
<td>1092 (61.1)</td>
<td></td>
</tr>
<tr>
<td>&gt;3 years</td>
<td>284 (26.3)</td>
<td>84 (12.4)</td>
<td>368 (21.1)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: HEW, health extension worker.

* Frequencies and weighted percentages reported.

**Factors Associated With Delayed Contraceptive Implant Removal in Ethiopia**

*Includes 2 respondents who said their implant was missing.

*Includes 92 respondents who still had their implant inserted at the time of interview.
calculated days past recommended removal ranged from 5 to 1,377 days. These women gave various reasons for not having had their implant removed, including they were planning on getting it removed soon (23%), did not know the removal date (22%), faced barriers to getting it removed (14%), and miscellaneous other reasons (41%) (results not shown).

In the multivariable logistic regression analyses, respondents who had their 1-rod contraceptive implant inserted by an HEW had 2.5 times the odds of keeping the implant for more than 3 years (adjusted odds ratio [aOR]=2.50; 95% confidence interval [CI]=1.19, 5.24) (Table 3). Living in SNNPR compared with Tigray (aOR=2.74; 95% CI=1.40, 5.35) or reporting distance or transportation to the facility as a barrier to removal (aOR=3.31; 95% CI=1.61, 6.78) were also factors associated with keeping the implant for longer than 3 years. Women who were literate/had any education (aOR=0.73; 95% CI=0.54, 1.00) and those who were married (aOR=0.55; 95% CI=0.39, 0.79) were significantly less likely to have kept their implant beyond 3 years. When sociodemographic characteristics were interacted with provider who inserted the implant (Table 4), older women and women of other religions who had their implant inserted by an HEW were significantly more likely to have kept their implant for more than 3 years, compared with women in those same categories whose implant was inserted by another health provider.

Survey respondents also reported if they had experienced any barriers to accessing implant removal. Table 5 shows that among women who had their implant removed at or before 3 years, those who had their implant inserted by an HEW were significantly more likely to report experiencing barriers, including that the provider was unable or refused to provide removal or that they faced challenges in transportation or the distance to the facility. Among the women who had kept their single-rod contraceptive implant for longer than 3 years, there were no significant differences by provider who inserted the contraceptive implant in barriers reported to implant removal, although about one-third of these women said they forgot or did not know their removal date or that the provider was unavailable when they visited the facility.

### DISCUSSION

Providing community-based health services in Ethiopia has helped the FMOH to significantly

---

**Table 3. Multivariate Logistic Regression Model for the Association Between Health Provider Who Inserted Implant and Keeping Implant for Longer Than 3 Years**

<table>
<thead>
<tr>
<th>Health extension worker (ref: other health worker)</th>
<th>2.50c (1.19, 5.24)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Client Age, years (ref: 25–34)</td>
<td></td>
</tr>
<tr>
<td>18–24</td>
<td>1.46 (0.58, 3.68)</td>
</tr>
<tr>
<td>35–44</td>
<td>1.20 (0.83, 1.72)</td>
</tr>
<tr>
<td>45–49</td>
<td>1.16 (0.58, 2.31)</td>
</tr>
<tr>
<td>Literate/any education (ref: illiterate/no education)</td>
<td>0.73b (0.54, 1.00)</td>
</tr>
<tr>
<td>Married (ref: unmarried)</td>
<td>0.55b (0.39, 0.79)</td>
</tr>
<tr>
<td>Orthodox (ref: any other religion)</td>
<td>1.07 (0.54, 2.12)</td>
</tr>
<tr>
<td>Region (ref: Tigray)</td>
<td></td>
</tr>
<tr>
<td>SNNPR</td>
<td>2.74c (1.40, 5.35)</td>
</tr>
<tr>
<td>Oromia</td>
<td>2.25 (0.86, 5.90)</td>
</tr>
<tr>
<td>Amhara</td>
<td>1.68 (0.64, 4.44)</td>
</tr>
<tr>
<td>Reported distance to facility/transportation as a barrier</td>
<td>3.31c (1.61, 6.78)</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; OR, odds ratio; SNNPR, Southern Nations, Nationalities, and Peoples’ Region.

a This model does not include any interaction terms.

b P<.05
c P<.001
Factors Associated With Delayed Contraceptive Implant Removal in Ethiopia

Getting to the removal date, or not knowing the removal date, or having transportation or distance barriers in returning to the facility. These barriers are reminiscent of some of the challenges to implant removal that have been documented in other contexts. For instance, a recent study in Ghana found that about half of women did not know that they could get their implant removed at a different health facility from the one where they had the implant inserted. After Implanon was introduced in South Africa, a study found that almost all women reported knowing that the implant should be removed after 3 years, but only two-

The most common barriers to removal among respondents who had their implant beyond 3 years were forgetting, not knowing the removal date, or having impediments getting to the facility.

### TABLE 4. Adjusted<sup>a</sup> Odds Ratios Comparing Health Extension Workers and Other Health Workers for the Association Between Exposure Variable and Keeping Implant for 3 Years or Longer (N=1,860)

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Adjusted OR&lt;sup&gt;b,c&lt;/sup&gt; (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>18–24</td>
<td>0.65 (0.17, 2.53)</td>
</tr>
<tr>
<td>25–34</td>
<td>2.45 (0.82, 7.36)</td>
</tr>
<tr>
<td>35–44</td>
<td>6.17&lt;sup&gt;d&lt;/sup&gt; (2.31, 16.48)</td>
</tr>
<tr>
<td>45–49</td>
<td>0.74 (0.10, 5.97)</td>
</tr>
<tr>
<td>Religion</td>
<td></td>
</tr>
<tr>
<td>Orthodox religion</td>
<td>0.87 (0.43, 1.78)</td>
</tr>
<tr>
<td>Any other religion</td>
<td>3.08&lt;sup&gt;e&lt;/sup&gt; (1.07, 8.82)</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; OR, odds ratio.
<sup>a</sup> Odds ratio adjusted for age, education, marital status, religion, region, and facility/transport barriers.
<sup>b</sup> Odds ratios shown for each subgroup to interpret significant interaction effects.
<sup>c</sup> Only covariates with significant interactions shown.
<sup>d</sup> P<.001.
<sup>e</sup> P<.05.

Accelerate scale-up of LARCs, and training HEWs in single-rod etonogestrel contraceptive implant insertion is major part of this effort. However, our analysis demonstrates that alongside this increase in access, Ethiopian women have faced challenges to obtaining timely removals. In particular, women who received their implants from HEWs were significantly more likely to keep their implants inserted beyond the recommended 3-year timeframe. Research on the experience of introducing contraceptive implants in South Africa suggests that providers may perceive that women discontinue implants at a higher rate than other methods and feel that early removal is “wasteful” and therefore discourage it.14 In this study, we did not find high proportions of women keeping implants due to provider refusals to remove them. However, clients of HEWs were more likely to report provider refusals, so it is possible that HEWs refused to provide removals because they were not trained to do so.

It is also possible that human resource capacity was insufficient to meet the demand for removal services among participants who had their implants inserted by HEWs. Although health center staff were trained in implant removals, each health center is a referral site for 5 health posts and the topography and infrastructure in Ethiopia often make travel difficult. This is likely to have particularly been the case in SNNPR, which is Ethiopia’s most rural region. As part of an effort to reach targets set in the National Reproductive Health Strategy to expand access to LARCs, the Ethiopian FMOH has recently introduced an initiative to train select HEWs in both the insertion and removal of implants and intrauterine devices.13 This effort is especially relevant in light of our findings, which demonstrate that even when implant introduction efforts try to build in safeguards to ensure timely removals (deploying higher-level providers to the community level and referring women to higher-level facilities for removals), they may not be sufficient to fulfill the removal needs of all women, particularly those living in rural areas who cite transportation or distance to the facility as barriers to care. Other projects piloting the “task-shifting” of implant provision have also demonstrated effective insertion skills by providers,14 but information on removal services available to women through these efforts is still limited.

Married women and women with more education were less likely to have kept their implant for more than 3 years. Though women were not asked about the reasons for having their implants removed, these differences could be due to married women being more likely to remove their implant to have another child and those with more education better comprehending when and at what time point they should have their implant removed. Interaction models indicated that older women and women of other religions who had their implants inserted by HEWs were significantly more likely to be using them past 3 years. For older women, they may have completed their childbearing and not see the necessity of seeking removal; however, it is unclear if these women believe that the implant may prevent pregnancy for longer than 3 years.

The most common barriers to removal among respondents who kept their implant beyond 3 years were that they forgot, did not know the removal date, or faced transportation or distance barriers in returning to the facility. These barriers are reminiscent of some of the challenges to implant removal that have been documented in other contexts. For instance, a recent study in Ghana found that about half of women did not know that they could get their implant removed at a different health facility from the one where they had the implant inserted.18 After Implanon was introduced in South Africa, a study found that almost all women reported knowing that the implant should be removed after 3 years, but only two-
thirds were told at insertion how long the implant can be used for and half were told when to return for removal. One woman specifically noted that she faced difficulties accessing removal services, first being told that the provider who originally inserted the implant was not available and being referred to another clinic, which subsequently told her she had to return to the same clinic where the implant was inserted for removal.\textsuperscript{19}

Recommendations from the Indonesia Norplant experience were that programs introducing implants should plan to have removal services available from the time that insertions are first provided.\textsuperscript{20}

Strengths of this study were its large sample size and rigorous sampling process, which was implemented to gain a sample generalizable to Ethiopian women’s experience with 1-rod etonogestrel contraceptive implant insertion and removal by HEWs and other health care providers across the 4 regions of the country where the implant scale-up initiative was implemented. We surveyed women who had had an implant inserted 3–6 years prior so that we could assess their experience obtaining removals through the full recommended life of the implant. In addition, because we identified participants for this study from the facility registers, we presumed that women in this subset had not moved since the time of insertion. One might expect that their knowledge of where to access removal should be greater than users who have since relocated or migrated. Therefore, it is notable that this study population of users still faced challenges.

Limitations

The study also had limitations, most notably challenges in obtaining accurate sampling frames of woredas that had started providing single-rod etonogestrel contraceptive implants at the community level, sampled sites that were inaccessible due to road conditions or a lack of eligible women that then required replacement, and security challenges in Amhara during the study period that curtailed data collection in that region and slightly reduced the precision of our estimates. As with any study collecting data retrospectively, participants’ responses may have been affected by recall bias. In this particular case, because Implanon is recommended to be used for up to 3 years, that

| TABLE 5. Barriers to Removal by Timing of Removal and by Provider That Inserted Implanon\textsuperscript{a} |
|-----------------|------------|------------|---------|----------|
| Removal at or before 3 years | HEW | Other Health Provider | Total | P-value |
| No barriers | n=1,062 | n=430 | n=1,492 | 1,236 (86.4) | .029 |
| Service provider unavailable day of visit | 47 (3.9) | 16 (3.7) | 63 (3.8) | .88 |
| Service provider unable to provide removal\textsuperscript{b} | 15 (1.8) | 4 (0.2) | 19 (1.1) | <.001 |
| Service provider refused to provide removal | 59 (5.7) | 21 (2.4) | 80 (4.3) | .015 |
| Distance to facility/transportation | 72 (4.8) | 12 (2.1) | 84 (3.6) | .020 |
| Other barriers\textsuperscript{c} | 20 (1.9) | 13 (0.9) | 33 (2.0) | .90 |
| Removal more than 3 years after insertion/still inserted | n=284 | n=84 | n=368 | 148 (37.8) | .61 |
| No barriers | 110 (38.7) | 38 (34.6) | 148 (37.8) | .67 |
| Forgot or did not know removal date | 89 (32.6) | 29 (31.4) | 118 (32.3) | .88 |
| Service provider unavailable day of visit | 80 (28.1) | 15 (25.6) | 95 (27.6) | .76 |
| Service provider unable to provide removal\textsuperscript{b} | 18 (5.8) | 5 (4.6) | 23 (5.6) | .39 |
| Service provider refused to provide removal | 16 (3.8) | 7 (6.2) | 23 (4.3) | .50 |
| Distance to facility/transportation | 43 (11.8) | 2 (9.9) | 45 (11.4) | .50 |
| Other barriers\textsuperscript{c} | 61 (26.7) | 23 (32.9) | 84 (28.1) | .50 |

Abbreviation: HEW, health extension worker.
\textsuperscript{a} Frequencies and weighted percentages reported. Respondents could select more than one barrier.
\textsuperscript{b} Includes lack of materials to provide removal.
\textsuperscript{c} Other barriers include lack of time to go to facility, fear, and cost.
may be the date that most women remembered, therefore we may be under- or overrepresenting the number of women who re-ported removing their implant within the 3 years duration. (Notably, respondents reporting using their implant for exactly 3 years constituted the largest proportion of our sample.) Finally, survey questions were limited to asking about any barriers to removal for each woman’s specific duration of use. Therefore, we were limited in our ability to assess a few relevant factors because they were not que-ried in the survey; these included asking whether women sought multiple removal attempts, if women who kept their implant beyond 3 years did so because they were continuing to rely on it for pregnancy prevention, and whether women who had their implants removed obtained removals through organized outreach activities at the health posts or at other health facilities.

CONCLUSION

Given the dramatic increase in implant use globally over the past decade, ensuring access to timely removals is a responsibility of family planning programs guided by a rights-based approach to voluntary use of contraception. Although clinical data suggest that the effectiveness and safety of 1-rod etonogestrel contraceptive implants persist beyond 3 years, if a woman wants an implant removed (to become pregnant or for any other reason), not being able to obtain removal on demand is problematic and in conflict with principles of agency, autonomy, and empowerment. Access to removal services in Ethiopia has likely already improved in the nearly 4 years since this study was conducted. However, given efforts to promote task sharing of implant services at lower levels of health care systems worldwide, the Ethiopian experience is still important to document and learn from. Future efforts to increase access to implants may need to include innovative, locally appropri-ate mechanisms to remind women of the duration for which they have had their implant inserted to alert them of the recommended timeframe for re-moval as well as removal training for all health care providers who provide implant insertions.

Acknowledgments: This study was conducted by the Federal Ministry of Health (FMoH) and FHI 360. In addition to the coauthors on this manuscript, we would like to acknowledge the roles of Girma Gemach, Family Planning Senior Expert at the FMoH, and the following PROGRESS/Ethiopia team members for their technical assistance with the design and implementation of this study: Admasu Terefe, Bethelhem Fekade, Bezawit Dagne, Wandimu Magrent, and Zenube Amin. We are further grateful to the study participants for their time, the health centers and health posts surveyed for granting permission for us to conduct this study at their health facilities, and the Regional Health Bureaus, Zonal Health Department, and Woreda Health Offices for Amhara, Oromia, Tigray, and SNNPR for their support for this study. Finally, we would like to thank the following FHI 360 headquarters staff for their assistance with data analysis and manuscript review: Tom Grey, Mario Chen, and Rebecca Callahan.

Funding: This work is made possible by the generous support of the American people through the United States Agency for International Development (USAID) agreement #663-14-000001. The content and views expressed by the authors in this publication do not necessarily reflect the views of USAID, the United States Government, or FHI 360.

Competing interests: None declared.

REFERENCES


Practical Implications of Policy Guidelines: A GIS Model of the Deployment of Community Health Volunteers in Madagascar

Aurélie Brunie,a James MacCarthy,b Brian Mulligan,c Yvette Ribaira,c Andry Rabemanantsoa,c Louisette Rahantanirina,c Caleb Parker,b Emily Keyesd

Key Findings

- Community health workers (CHWs) in 23% of communities are assigned to cover either more than 1,000 people or more than 25 km².
- In 58% of communities, CHWs were more than 2 hours away one way from the health facility; in 61% of communities, CHWs were more than 2 hours from their assigned supply point.

Key Implications

- Program managers should consider assigning CHWs to the nearest health facility and placing supply facilities closer to decrease travel time.
- Policy makers should give latitude to communities and health centers to select and manage a number of volunteers matching local geographic and population characteristics.

ABSTRACT

Background: With increasing interest in strengthening community health programs nationally comes a need for operationalizing them in a realistic and achievable way. Limited information is available to help program managers establish appropriate parameters for their context. We examined aspects of program implementation related to deployment patterns of community health workers, called agents communautaires or ACs, in 2 districts of Madagascar.

Methods: By analyzing program data and publicly available datasets in a geographic information system (GIS), we estimated the population and surface area coverage expected of ACs in 445 fokontany (communities). Additional modeling on travel time demands examined 1-way pedestrian travel time for ACs to receive routine support from their assigned health facilities and from socially marketed supply points under dry season conditions, as well as the impact on travel time based on ACs being reassigned to other facilities or supply points.

Results: With the current distribution, ACs in 90% of fokontany have a catchment population of 1,000 or fewer people (2020 estimates) and ACs in 84% of fokontany have a catchment area of 25 km² or less. We estimated that ACs in 58% of fokontany were located more than 2 hours from their supporting health facility, and the proportion of fokontany with ACs more than 2 hours away from their assigned supply point was 61%. Reassigning ACs to the closest facility or supply point led to modest improvements in those figures (7 and 4 percentage points, respectively).

Conclusion: Findings allow visualizing the practical implications of coverage ratios for ACs to assess whether current demands are realistic. The physical access between ACs and the health system warrants significant attention due to challenges in transport and logistics. Analyses are timely to inform the Ministry of Public Health’s strategic thinking in the context of the development of the National Strategic Plan on Strengthening Community Health.

INTRODUCTION

Community health workers (CHWs) have been a vital component of primary health care since their inception close to 50 years ago. Although there are some examples of well-run CHW programs at scale in countries such as Brazil, Bangladesh, or Nepal, challenges have also been documented in scaling up and maintaining CHW programs, including poor planning, lack of coordination among actors, donor-driven management,
A GIS Model of Community Health Volunteers in Madagascar

We conducted 2 sets of analyses that are relevant in the context of Madagascar and suited to the use of GIS. The first set of analyses pertains to how many CHWs to deploy relative to population size and to the size of their catchment area. Establishing a realistic target population size typically depends on a constellation of factors including expected workload, frequency of contact required, services provided, expected time commitment, and local geography. Currently, some variability exists across countries in how coverage ratios are specified, with assignments for example being defined as a number of people or households per facility or per village. There is not always a clear understanding of how different metrics relate to each other or how they translate in terms of the geographic area CHWs may need to move around depending on their assigned tasks.

The second set of analyses examines travel time demands associated with maintaining functional linkages to the health system. In settings where CHWs are required to travel to health facilities for supervision and logistical support and supplies, they can incur additional transport and opportunity costs. Although reducing the costs for the populations CHWs serve is recognized as a significant advantage of CHW programs, costs borne by the CHWs are not always adequately considered as part of program design, planning, and implementation. A reality check to verify that assignment patterns of CHWs to health facilities are rational and travel time demands realistic based on health facility location, distance, terrain, and road system quality is important to ensure that programs are well-functioning and sustainable.

**Methods**

**Setting**

The analyses examined the realities that volunteer CHWs face in performing their tasks in 2 districts of Madagascar (Mandritsara in Sofia region and Mananara Nord in Analanjirofo region) supported by the United States Agency for International Development (USAID) Community Capacity for Health Program, locally known as Mahefa Miaraka, implemented by JSI Research & Training Institute, Inc. Conducting this activity in Madagascar is timely because, after updating its National Community Health Policy in 2017, the Ministry of Public Health (MSANP) is currently developing the associated strategic plan. The main cadre of CHWs in Madagascar consists of volunteers called agents communautaires (ACs). Available estimates place the number of trained ACs at over 34,000 across the
community to their travel time by foot as well as their expected of ACs, coverage geographic population and We examined the GIS Model of Community Health Volunteers in Madagascar www.ghspjournal.org research. Ethics and deemed to be exempt from ethical approval because it was not human subjects by FHI 360. Through the MSANP, we obtained a database with the full list of ACs in the 2 districts, their fokontany and their assigned CSB, along with the geo-coordinates of the CSB. We downloaded administrative boundaries for regions, districts, commune (administrative collection of fokontany), and fokontany in Madagascar from the United Nations Office for the Coordination of Humanitarian Affairs data portal, and manually matched fokontany names from the MSANP list to administrative boundary shape files. We estimated fokontany population by aggregating 2020 estimates of population per 100 by 100-m grid-cell produced by WorldPop. Fokontany surface areas were derived within the GIS using standard tools. Using information on the number of ACs assigned to each fokontany from the MSANP database, we calculated the population-to-AC and the surface-area-to-AC ratios for each fokontany. Where there were multiple ACs within the same fokontany, we assumed that population and surface area were split evenly among them.

GIS Analysis Approach
Using multiple data sources and modeling within a GIS, we conducted 2 sets of analyses. In the first set, we examined the population and geographic coverage expected of ACs, as defined by the population-to-AC ratio and the surface-area-to-AC ratio. In the second set, we modeled the 1-way pedestrian travel time from the fokontany that the ACs serve to their assigned CSB under dry season conditions, as well as the 1-way pedestrian travel time to ACs’ assigned PA. Finally, we modeled the impact on travel time of scenarios whereby ACs were reassigned from their current CSB or PA to their closest CSB or PA. We used R version 3.5.1 for cleaning and plotting data, and ArcGIS Pro version 2.4.0 for all spatial analyses and map creation. This activity was reviewed by FHI 360’s Office of International Research Ethics and deemed to be exempt from ethical approval because it was not human subjects research.

Expected Coverage Analysis
Through the MSANP, we obtained a database with the full list of ACs in the 2 districts, their fokontany and their assigned CSB, along with the geo-coordinates of the CSB. We downloaded administrative boundaries for regions, districts,
considered this assumption to be reasonable since PAs are typically co-located with small businesses and found in urban and peri-urban areas. Fourteen fokontany were excluded from analyses on travel time to PA because there was no PA identified for the corresponding commune. Because we did not have information about where each AC was located, we approximated the location for ACs based on potentially habitable land in each fokontany using spatially disaggregated population data from WorldPop. We considered any area with 10 people or more per square kilometer to be habitable and a potential AC location. We chose a conservative population density threshold that contained more than 93% of the population across the 2 districts to ensure that we considered travel time requirements for ACs that live in urban and rural locations.

**Cost Distance Raster**

Since opportunities for mechanized travel are limited in the 2 districts, we used a cost distance raster (i.e., a grid of cells containing values reflecting the amount of time it takes to travel across each cell) to model travel time across roads and all terrain. We assumed travel by foot, as is typical for ACs in the 2 districts. To create the cost distance raster, each 100 by 100-m square cell was assigned a value representing the “cost” of traveling across the cell, expressed as a total time in minutes. The travel time cost for each cell was calculated by combining several layers of data to represent local conditions and assigning varied travel speeds corresponding to these conditions.

Local conditions were represented by the availability and types of roads, elevation, land cover, and rivers. The road network dataset was downloaded from OpenStreetMap after digitizing 885 km of roads and paths that were missing from the database for the 2 districts.26 The source of the geographic elevation was the 30-m resolution SRTM Digital Elevation Model (DEM), downloaded from the United States Geological Survey EarthExplorer portal, and land use data at 300-m resolution came from the European Space Agency’s GlobCover project.27,28

We used a baseline walking speed of 5 km/h for bare areas and all road types and reduced it for other land cover types (e.g., dense forests) to represent reduced walking speeds as in the GlobCover data.29,30 We used elevation data to model the effect of slope on travel speeds using Tobler’s hiking function and to generate a river network using the D8 flow accumulation method.17,31 We then used the Strahler stream order method to approximate the width of each river and defined river crossing speeds based on estimated widths and possible delays from waiting for a dugout canoe for medium rivers (20–60 m wide) or ferry for large rivers (over 60 m wide).32 Estimates of walking speed and river crossing delays were validated with Mahéfia Miaraka field teams.

**Travel Time**

We used the cost distance raster to model 1-way fokontany-level pedestrian travel times by averaging travel times between potential AC locations within each fokontany and the CSB or PA assigned to the ACs from that fokontany. We then similarly modeled fokontany-level travel times to the closest CSB or PA. CSBs that did not have any ACs assigned to them in the MSANP database were not considered suitable for reassignment of ACs. All estimates represent travel times under dry season conditions and assume that each habitable cell in the raster has an equal probability of containing an AC.

**RESULTS**

**Expected Coverage**

Table 1 shows characteristics of the 2 districts. Overall, 962 ACs were deployed across 445 different fokontany and supported by a network of 56 CSBs and 47 PAs. Three other CSBs did not support any ACs. The number of ACs per fokontany ranged from 1 to 6 (Figure 1); 92% of fokontany had 2 ACs, 7% had between 3 and 6, and the remaining 5 fokontany had fewer than 2. Implementing partners support the selection of 2 volunteers per fokontany as a matter of course. Additional volunteers can be added at the discretion of the CSB. Fokontany with more than 2 ACs most likely respond to greater coverage needs that arise from a larger population or surface area, a more dispersed population, geographic barriers, and/or different population subgroups (e.g., ethnic groups). Fokontany with less than 2 ACs likely are explained by 1 AC having recently stepped down and not yet been replaced, or where few people meet selection criteria.

Overall, 89% of fokontany had an estimated 2020 population of 2,000 or fewer people (range: 145–11,359) and 95% of fokontany spanned 50 km² or less, with the largest fokontany covering 365.4 km² (range: 0.1–365.4 km²). The population-to-AC ratio was at or below 1,000 in 90% of fokontany, and all but one fokontany had an expected population
coverage of 2,000 people or fewer per AC (Figure 2). The surface-area-to-AC ratio was 25 km² or less in 84% of fokontany, with the expected surface area coverage exceeding 50 km² per AC in 5% of fokontany (Figure 3). One AC had an assigned area of 363 km².

Figure 4 presents the overlay of population and surface area coverage. When combining the 2 measures, we found that 77% of fokontany in which ACs were assigned had 1,000 people or fewer and an area of 25 km² or less.

**TABLE 1.** Characteristics of 2 Districts in Madagascar and Distribution of Community Health Workers

<table>
<thead>
<tr>
<th>District</th>
<th>Mandritsara</th>
<th>Mananara Nord</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region</td>
<td>Sofia</td>
<td>Analanjirofo</td>
</tr>
<tr>
<td>Population</td>
<td>323,242</td>
<td>216,281</td>
</tr>
<tr>
<td>Number of fokontany</td>
<td>239</td>
<td>206</td>
</tr>
<tr>
<td>Number of CSB with ACs</td>
<td>35</td>
<td>19</td>
</tr>
<tr>
<td>Number of PA</td>
<td>30</td>
<td>17</td>
</tr>
<tr>
<td>Number of AC</td>
<td>530</td>
<td>432</td>
</tr>
</tbody>
</table>

Abbreviations: AC, agents communautaires (community health workers); CSB, centres de santé de base (health centers); PA, points d’approvisionnement (supply points).

**FIGURE 1.** Number of ACs per Fokontany in Madagascar

Abbreviation: ACs, agents communautaires (community health workers).

**Travel Time**

Figure 5 and 6 show the estimated 1-way travel times to ACs’ assigned CSB and PA, respectively, by foot and under dry season conditions. ACs were within 2 hours by foot of their assigned CSB in 42% of fokontany, whereas travel time to the CSB was between 2 and 4 hours in 34% of fokontany and greater than 4 hours in 24% of fokontany. Travel time to the assigned PA was 2 hours or less in 39% of fokontany, between 2 and 4 hours in 33% of fokontany, and greater than 4 hours in...
the remaining 27%. The median travel time to the CSB and PA was similar (2.3 and 2.5 hours). In both cases, travel time varied greatly across fokontany, ranging from 3 minutes to more than 10.5 hours for CSBs and from 4 minutes to over 11 hours for PAs (Table 2).

Overall, ACs in 31% of fokontany were not assigned to their closest CSB, and ACs in 25% of

FIGURE 2. Population Coverage per ACs in Madagascar

Abbreviation: ACs, agents communautaires (community health workers).

FIGURE 3. Surface Area to ACs in Madagascar

Abbreviation: ACs, agents communautaires (community health workers).
fokontany were not assigned to their closest PA. After reassigning all ACs from their assigned to their closest CSB, the mean travel time decreased from 2.86 hours to 2.39 hours. The proportion of fokontany with ACs within 2 hours of travel time to the CSB increased from 42% to 49% and that of fokontany with travel time between 2 and 4 hours from 34% to 38%, and the proportion of fokontany

**FIGURE 4.** Number of Fokontany per Population and Surface Area Covered by ACs

Abbreviation: ACs, agents communautaires (community health workers).

**FIGURE 5.** Estimated Travel Time for AC From Each Fokontany to Their Assigned CSB by Foot During the Dry Season (1 Way)

Abbreviations: ACs, agents communautaires (community health workers); CSB, centres de santé de base (health centers).
where ACs were confronted with more than 4 hours or travel decreased from 24% to 13%.

Reassignment from the assigned to the closest PA led to a reduction in mean travel time from 3.11 hours to 2.74 hours. Similar to CSBs, there was an increase in the proportion of fokontany with a travel time of 2 hours or less (39% to 44%) or between 2 and 4 hours (33% to 35%) and a decrease in the proportion of fokontany with the least favorable travel time, over 4 hours (27% to 21%).

### DISCUSSION

We used GIS to support program managers in visualizing data patterns related to the current deployment patterns of ACs in 2 districts of Madagascar. Our focus was primarily on improving...
These analyses are meant to provide managers with information for assessing whether the status quo is rational and realistic and developing plans for adaptive management.

Our results offer an important first step in contextualizing guidelines and assessing whether resulting demands on ACs are inherently realistic.

Under the Mahefa Miaraka program, ACs are allowed to sell at a small markup to the actual cost of health commodities and they also receive a modest allowance and a travel reimbursement (based on MSANP guidelines) for attending meetings that may not be sufficient to offset opportunity costs and possible expenditures, particularly for ACs who travel the farthest distances. Furthermore, with a Rural Access Index of 11.4%, an estimated 17 million people in rural areas of Madagascar have a catchment area of 25 km² or less. In 2020 estimates, ACs in 84% of fokontany have a catchment population of 1,000 people or fewer. In 445 fokontany in the 2 districts largely aligned with the guideline in use of 2 ACs per fokontany that is supported by implementing partners. With the current distribution, ACs in 90% of fokontany have a catchment population of 1,000 people or fewer (2020 estimates) and ACs in 84% of fokontany have a catchment area of 25 km² or less. In 23% of fokontany, ACs are assigned either more than 1,000 people or more than 25 km². These results only provide rough estimates given several assumptions made about the specific location of ACs and population patterns within fokontany. ACs who travel the farthest distances. Further, the thresholds we used (1,000 persons and 25 km²) are based on descriptive patterns observed in our data rather than an informed determination of what may be manageable. Experience from other countries and the WHO guideline on CHW programs indicate that there is no single, ideal target population size per CHW but rather that the optimal coverage ratio depends on a constellation of factors, including local epidemiology, CHWs scope of practice, geographic distribution of the population, geographic accessibility, and the balance between time demands and compensation and incentives. Thus, policy makers need to take these other factors into consideration alongside our findings and allow greater latitude for local communities and CSBs to reasonably select and manage the number of volunteers based on their circumstances. Our results offer an important first step in contextualizing guidelines and assessing whether resulting demands on ACs are inherently realistic.
Madagascar are unconnected to the road network (i.e., do not live within 2 km of the nearest road in good condition). Thus, the link between ACs and the health system warrants significant attention and may require different solutions beyond the reassignment explored in our analyses.

Although endowing ACs with bicycles, as has been done in some contexts, may cut travel times, this solution is not currently considered under the National Strategic Plan for Strengthening Community Health, largely because ACs are supported solely through partner-supported programs. In addition, bicycles would be difficult to use in the program area because of the sandy, muddy terrain during the rainy season. One option that is currently being discussed by the MSANP and implementing partners is to switch to a single-source resupply model exclusively through health facilities, suppressing the need for additional trips to supply points. Complementary measures ensuring that a monthly trip to the health center can serve as a 1-stop shop for supervision, reporting, and resupply to minimize travel burden should be considered. These may include commodity security at the health facility level, strong commodity management skills at the AC level, reliable communication systems to avoid unnecessary trips when supplies are lacking or staff is not available, and AC incentives. Transport and logistics challenges are compounded during the rainy season when some areas are entirely cut off, largely due to swelling of rivers. Rainfall patterns vary across Madagascar and have been disrupted due to changes in climate, with northern Madagascar receiving 150% of expected annual rainfall in 2018. Adaptive management that considers the dynamic impact of the rainy season on access will be increasingly important to reduce service disruptions and increase health system resilience. Additional creative interventions may be required for reliable access to communities, such as using hovercrafts for deliveries as was recently tested in another program.

### Limitations

Analyses were conducted in 2 districts of Madagascar supported by the same community health program and may not adequately represent other areas of the country. As in any modeling exercise, our results are anchored in a combination of hard data and assumptions, for example, about the precise location of ACs within each fokontany. Obtaining geocoordinates for each AC was cost prohibitive and not feasible within the timeline of the project. Although our modeling assumptions were grounded in discussions with local staff, findings should be interpreted cautiously as directional rather than as a precise representation of the actual conditions that exist in the 2 districts. Population and geographic coverage are unlikely to be equally divided between ACs within the same fokontany. One of the most critical assumptions related to estimation of delays with river crossings is limited by the availability of data on river characteristics. In addition, modeling was performed under dry season conditions; a similar approach could be used with rainy season parameters and would likely lead to increased travel times. When interpreting findings, program managers should also consider how differences in the profile of ACs, such as those based on gender, age, occupation, or family responsibilities may affect reasonable expectations in terms of the demands placed on ACs.

### Conclusion

A major contribution of CHW programs is to bridge the gap between communities and formal health services. An earlier study in another area of Madagascar showed that CHWs share in the living conditions of the populations they serve. Complementing these findings, our analyses highlight that, although CHW programs reduce challenges for the client population to access an integrated package of services for women and children, access barriers are in fact transferred onto CHWs.

Using GIS to visualize the deployment patterns of CHWs can improve program managers’ ability to synthesize information and grasp the actual implications of policy decisions, and modeling within a GIS enables identification of data patterns related to the demands placed on CHWs, providing useful information to inform decision making. This information is timely to inform the MSANP’s strategic thinking around criteria and processes for optimal integration of ACs into communities and the health system in the context of the development of the National Strategic Plan on Strengthening Community Health. Our findings suggest that policy makers should consider allowing greater latitude to reasonably select and manage the number of ACs that match the realities of each community. In addition, program managers should consider assignment of ACs and placement of supply facilities that decrease travel time.

More broadly, optimization will require additional research to more fully understand the realities of ACs, and management and real-time data use to assess what is required to ensure functional
support systems. Specific recommendations to balance these considerations include the following:

1. Conduct additional geospatial analysis including a larger sample of geographic areas to obtain a broader representation of the local realities of ACs in the country. Incorporate additional analysis of existing data from the HMIS to examine how coverage and travel time may correlate with functioning of community health systems.

2. Use task analysis and time-use research to better understand how ACs manage their tasks and whether expectations and workloads are realistic and commensurate with compensation and incentives.

3. Use costing tools to provide an accompanying estimate of the real costs of maintaining a volunteer community health system and required inputs.

4. Review management data on health staffing at CSBs to assess the ability of the health system to effectively work with groups of ACs, including for supervision and to ensure the quality of reporting, review, restock of supplies and continued learnings during monthly meetings.

Acknowledgments: We acknowledge Alexis Hoyt, Victoria Lebrun, and Patrick Olsen for assistance with the digitization of the road network. We are grateful to Dr. Jean-Claude Andriamirandrazimisoa, Secretary General, and the Ministry of Public Health for their support to this activity and for granting us access to the necessary data for this exercise. We are also thankful to USAID/Madagascar, including Ms. Azzah Al-Rashid, Dr. Jocelyne Andriamisonanana, and Dr. Andry Rahajarison, for their guidance and oversight.

Funding: This activity was completed with funding from the United States Agency for International Development (USAID) Mission in Madagascar under the terms of Cooperative Agreement AID-OAA-A-12-0004, the Advancing Partners and Communities (APC) project. The information contained in this document is the sole responsibility of FH 360 and does not necessarily reflect the position of USAID or partner organizations.

Competing interests: None declared.

REFERENCES


Private Providers’ Experiences Implementing a Package of Interventions to Improve Quality of Care in Kenya: Findings From a Qualitative Evaluation

Masila Syengo, Lauren Suchman

Key Findings

- Engaging private providers in comprehensive quality improvement activities is achievable.
- Cost is a prohibiting factor for many private providers who would like to implement quality improvement interventions.
- Offering a package of quality improvement interventions may help mitigate some cost issues as opposed to a one-size-fits-all approach.

Key Implications

- Program managers should consider tailoring several quality improvement interventions to meet the needs of private providers in small and medium-sized facilities.
- Researchers should consider conducting further studies on the return on investment quality improvement programs can offer to private providers in LMICs.

ABSTRACT

Introduction: Quality of care is an important element in health care service delivery in low- and middle-income countries. Innovative strategies are critical to ensure that private providers implement quality of care interventions. We explored private providers’ experiences implementing a package of interventions intended to improve the quality of care in small and medium-sized private health facilities in Kenya.

Methods: Data were collected as part of the qualitative evaluation of the African Health Markets for Equity (AHME) program in Kenya between June and July 2018. Private providers were purposively selected from 2 social franchise networks participating in AHME: the Amua network run by Marie Stopes Kenya and the Tunza network run by Population Services Kenya. Individual interviews (N=47) were conducted with providers to learn about their experiences with a package of interventions that included social franchising, SafeCare (a quality improvement program), National Hospital Insurance Fund (NHIF) accreditation assistance, and business support.

Results: Private providers felt they benefited from trainings in clinical methods and quality improvement offered through AHME. Providers especially appreciated the mentorship and guidelines offered through programs like social franchising and SafeCare, and those who received support for NHIF accreditation felt they were able to offer higher quality services after going through this process. However, quality improvement was sometimes prohibitively expensive for private providers in smaller facilities that already realize relatively low revenue and the NHIF accreditation process was difficult to navigate without the help of the AHME partners due to complexity and a lack of transparency.

Conclusion: Our findings suggest that engaging private providers in a comprehensive package of quality improvement activities is achievable and may be preferable to a simpler program. However, further research that looks at the implications for cost and return on investment is required.

INTRODUCTION

Quality of care is an important component of Sustainable Development Goal 3, which aims to promote healthy lives and well-being for all ages through improved quality measures.1,2 However, as more low- and middle-income countries (LMICs) attempt to achieve universal health coverage,3 it will be important not to lose sight of quality in the race to reach a larger number of people. Indeed, we know that
quality is just as important as quantity when it comes to achieving meaningful health outcomes across an entire population.4,5

Studies have shown that improving the quality of care is marred by complex factors, including health workers multitasking, lack of training, underuse of patient management protocols, weak supportive supervision in an environment of health worker staff shortages, and weak policy initiatives that tackle low quality of care implementation in high-poverty areas.6-8 Researchers have cited weak health system structural factors, such as longer waiting times and commodity stock-outs, among other compounded health system challenges affecting the quality of care in LMICs.9 In terms of maternal, neonatal, and child health care, many women in LMICs lack comprehensive quality of care throughout pregnancy and delivery.10 In Kenya, population-based research has shown that the poorest women received fewer essential services during ANC care and were 4 times more likely to deliver without a skilled attendant compared to those women in the wealthiest quintile.11

Tunçalp5 has shown that both clinical and non-clinical interventions implemented in LMICs can improve quality of care and in turn increase desired individual and facility-level health outcomes, as well as people-centered outcomes. Further, a positive relationship between patients and providers that is marked by good rapport, empathic communication, active listening, and confidentiality, was reported to have increased utilization of health services and improved quality of care and health outcomes in LMICs.9 Overall, improving quality of care increases the opportunity for patients’ and facility-level outcomes with a focus on maximizing the utilization of health care services.

Although several evidence-based strategies exist to achieve quality improvement in clinical settings,12 the African Health Markets for Equity (AHME) program partners specifically focused on social franchising, SafeCare, and accreditation with national health insurance to help private providers improve quality. Research has shown that social franchising has become an increasingly popular health system strengthening strategy in poor and underserved communities because of its ability to maximize the potential of the private sector and improve access to health care services.13 Social franchises are comprised of a network of members who are private health care providers that use a commercial branding identity to achieve a social cause rather than a financial goal.13 Private providers are organized in a contractual obligation to offer specific services within a specific network of providers. These franchisees are then provided training, branding, and monitoring with the aim of improving quality of care, increasing access to care, expanding the affordability of services, and rapidly increasing the number of delivery points for important public health services.14,15 Results from a systematic review led by Beyeler16 found an association between social franchising and increases in both client volume and satisfaction. However, it was not clear that social franchising increased health care utilization or health impact, and social franchise clinics tended to underperform in terms of cost effectiveness and equity in relation to their nonfranchised counterparts. Still, there is evidence to suggest that franchise providers are able to maintain support for their operations through patient-user fees, which allows them to offer quality services.17

SafeCare, a step-by-step holistic quality improvement strategy, differs from other quality improvement models that tend to target specific programs or services within a health care facility. SafeCare aims to improve patient safety and quality of health care across an entire facility by offering assessments and improvement strategies for all aspects of the clinic, ranging from administration and management to record keeping, inventory management, drug safety, and clinical infrastructure.18 Using this strategy, facilities with severe shortages in equipment, infrastructure, and resources are supported to achieve stepwise improvements focusing on the most important areas of quality, safety, and risk, which would have hindered service delivery and quality of care.19 Some evidence suggests that this approach makes both the patients and providers increase trust in the way in which health care provision is administered more transparently, thus increasing access to health.20 Further, existing reviews of SafeCare in LMICs have shown that the program improved access to quality of care in poor underserved populations in Ghana, Kenya, Nigeria, Tanzania, and Uganda.21 On the provider side, private providers implementing SafeCare can increase financial investment through quality improvement targets and access to credit. This increases motivation to improve the quality of services, thereby also potentially increasing the number of patients served, as well as facilities’ efficiency.22

Despite successes in the implementation of SafeCare, difficulties remain with financial sustainability being the main challenge. Providers are constrained and often unable to absorb the increased costs of the SafeCare improvements, even if higher quality of care is achieved.22 A
We sought to understand private providers’ experiences and challenges while participating in this quality improvement intervention and learn about other ways to improve health care quality among these providers in Kenya.

Finally, research on health insurance schemes has shown evidence of improved patient access and utilization of health care services, as well as quality improvement outcomes. For example, several studies have shown that being insured was associated with adherence to treatment. On the private providers’ side, health insurance accreditation has been shown to improve quality of care services in the United States, and there is consistent evidence that shows that accreditation processes improved care and clinical outcomes across a wide spectrum of clinical conditions in high-income countries. However, some studies have found mixed effects of health insurance. For example, Suchman found that national health insurance accreditation in Kenya and Ghana did seem to help private providers increase their quality of care, but it was unclear to what extent access to this quality care became more equitable in the face of the many challenges that providers faced navigating and being paid by insurance. Further, a study in South Africa showed that improved compliance with accreditation standards had minimal or no effect on clinical outcomes and another study identified weak or inconsistent relationships between accreditation and quality measures outcomes. These findings point to the conclusion that different quality measures should not be expected to promote similar outcomes.

It is worth noting that limited literature exists that explores quality improvement among private sector providers in Kenya, particularly those providers that operate small and medium-sized facilities. In this article, we seek to fill this gap by exploring private providers’ experiences with a comprehensive package of interventions meant to improve both quality and accessibility.

With an eye to promoting health care quality among a larger proportion of providers, this article examines private providers’ experiences with a package of interventions intended to improve the quality of small and medium-sized private health facilities in Kenya. Specifically, we seek to better understand private providers’ experiences in this context to determine whether providers felt their clinical quality improved through participation in a comprehensive package of quality improvement interventions, the challenges they faced, and what other opportunities might exist for improving health care quality in Kenya, particularly among private providers in smaller facilities.

METHODS

This article draws from qualitative data collected at 47 private facilities that offer comprehensive maternal and reproductive health care across Kenya. The majority of these providers were already participating in the African Health Markets for Equity (AHME) program in Kenya, which operated from 2012 to 2019. AHME sought to increase access to high-quality primary health care for low-income clients in Kenya through a comprehensive package of quality improvement interventions. The AHME partnership included Marie Stopes International, Population Services International, Population Services Kenya, and the PharmAccess Foundation. Past partners included the International Finance Corporation and the Grameen Foundation. The participating franchise networks in Kenya included the Amua franchise operated by Marie Stopes Kenya and the Tunza franchise operated by Population Services Kenya.

The AHME intervention package aimed to address 5 conditions intended to increase health market accessibility for poor populations: (1) primary health care is covered by national health insurance; (2) poor populations are enrolled into national health insurance; (3) private providers are accredited with insurance; (4) private providers offer quality services; and (5) private providers are able to run sustainable businesses. As such, the intervention package was meant to be comprehensive, addressing all 5 market conditions. The package included social franchising to: (1) organize private health care providers into networks to deliver a specific package of health services (in this case, family planning and maternal and child health) under a common brand aimed at delivering comprehensive care with a social mission; and (2) train providers in both clinical practice, basic facility management, and monitoring and oversight of clinical quality. SafeCare was included to address quality improvement.
Because quality improvement efforts can be costly for small private facilities that generate little revenue on top of their operating expenses, AHME also offered a business support intervention that included access to the Medical Credit Fund (MCF), a program that connects providers with banks that can offer them accessible loans at relatively low interest rates. The business support intervention also included general support for clinic financial management, such as training in bookkeeping. Finally, AHME-supported providers that were not already accredited with Kenya’s national health insurance scheme, the National Hospital Insurance Fund (NHIF), were given accreditation support, with AHME representatives offering pre-inspection checklists, walking providers through the application process and liaising with local NHIF offices on providers’ behalf.

In this article, we focus on data collected through the AHME qualitative evaluation, which was an external program evaluation conducted by the University of California San Francisco (UCSF). While the AHME qualitative evaluation was far-reaching and spanned the duration of the program, we narrow in on data collected in the final round of data collection (2018) with private providers regarding their experiences with the AHME quality improvement interventions in Kenya. We used a qualitative approach and conducted semistructured interviews to address the research objectives; a total of 47 individual interviews were conducted.

### Study Setting and Selection of Participants

The final round of qualitative data collection for the AHME evaluation in Kenya was conducted between June and July 2018 with providers in both AHME franchised facilities and nonfranchised facilities across 23 counties in the 6 regions in Kenya (Nairobi, Eastern, Coast, Central, Rift Valley, and Western) (Table).

The study team received lists of franchise providers from both the Amua and Tunza networks and used a purposeful sampling design to select providers according to their location, their level of participation in the different components of the AHME intervention package, and their NHIF accreditation status. The franchise networks also provided lists of facilities that had been approached to join a franchise network but had declined participation. Although these “matched” facilities were intended to serve as a comparison group to help the study team determine the effects of the AHME interventions, interviews with nonfranchised providers ultimately yielded little useful data. However, we have included the perspectives of some of these providers regarding accreditation with and participation in the NHIF to help illustrate the challenges that private providers face working with the NHIF regardless of their franchise status.

### Study Procedures

Interview guides were developed by the study team at UCSF and data collection was supervised by Innovations for Poverty Action (IPA), a research organization based in New Haven, CT, USA, with country offices in cities across the globe, including Nairobi. Field staff were hired by IPA to conduct the interviews, and staff were jointly trained by UCSF and IPA in qualitative interviewing techniques.

### TABLE. Characteristics of Private Providers Interviewed for Quality of Care Interventions, Kenya

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No. (N=47)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, median (range), y</td>
<td>44 (38–60)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>35</td>
</tr>
<tr>
<td>Female</td>
<td>12</td>
</tr>
<tr>
<td>Education</td>
<td></td>
</tr>
<tr>
<td>College/ diploma/ certificate</td>
<td>36</td>
</tr>
<tr>
<td>University</td>
<td>8</td>
</tr>
<tr>
<td>Masters/ doctorate</td>
<td>3</td>
</tr>
<tr>
<td>Facility Type</td>
<td></td>
</tr>
<tr>
<td>Hospital</td>
<td>6</td>
</tr>
<tr>
<td>Health center</td>
<td>13</td>
</tr>
<tr>
<td>Clinic</td>
<td>20</td>
</tr>
<tr>
<td>Maternity home</td>
<td>3</td>
</tr>
<tr>
<td>Dispensary</td>
<td>3</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
</tr>
<tr>
<td>Professional qualification</td>
<td></td>
</tr>
<tr>
<td>Medical doctor</td>
<td>3</td>
</tr>
<tr>
<td>Nurse</td>
<td>15</td>
</tr>
<tr>
<td>Community auxiliary nurse</td>
<td>21</td>
</tr>
<tr>
<td>Clinical officer</td>
<td>1</td>
</tr>
<tr>
<td>Lab tech</td>
<td>2</td>
</tr>
<tr>
<td>Admin and management</td>
<td>5</td>
</tr>
<tr>
<td>NHIF accredited</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>21</td>
</tr>
<tr>
<td>No</td>
<td>26</td>
</tr>
</tbody>
</table>
ethical research practices, and fieldwork protocols. After their training, the interviewers held mock interview sessions with clinic staff at franchised facilities in Nairobi; these facilities were then excluded from participating in the formal study. Staff from both IPA and UCSF supervised the pilot testing, which helped to enhance the study tools. Semi-structured interviews were then conducted with private providers who were participating in the AHME interventions and matched facilities that were within the catchment area of AHME and were not participating in the intervention package. Interviews lasted approximately 40–60 minutes and explored providers’ experiences with: social franchising, the NHIF application process and experiences with NHIF once they were accredited, the SafeCare program, and MCF. Specifically, providers were asked questions related to the benefits of each program and any challenges they faced implementing these programs in their own health facilities.

Participants’ sociodemographic information was collected after completing the interviews. Interviews were audio-recorded, and detailed field notes were collected. To ensure participant confidentiality, the study team received informed consent to participate from all the participants, and all interviews were conducted in a private space (e.g., a provider’s office or exam room) within the health facility. The field research team was comprised of all Kenyans who were also native Swahili speakers. The team conducted data collection in English or Swahili, based on the interviewee’s preference. Recordings were translated and transcribed by a team of professional transcriptionists who were also natives of Kenya and had been trained by IPA research staff. Transcripts were de-identified to further protect participant privacy and confidentiality. IPA research staff in Kenya were responsible for back-checking interviews, including ensuring translation accuracy. Participants were compensated with a small token of appreciation in recognition of their time taken to participate in the research. This was usually a bar of soap that was worth about 200 Kenyan Shillings (US$2).

Data Analysis
Following transcription, qualitative researchers at UCSF independently reviewed the transcripts and developed a coding framework using an applied iterative approach, with codes developed and adapted from earlier rounds of data analysis and configured along lines of significant inquiry. The UCSF and IPA team then reviewed the initial codebook together to ensure a common understanding of codes and consistency in code application. Codes were refined throughout the coding process to allow for emerging themes and new priorities in the analysis were verified for consensus between the qualitative researchers to ensure inter-rater reliability. After coding, emerging themes were organized according to barriers and opportunities to implementing the full package of quality improvement interventions. A standard qualitative analysis software package (ATLAS.ti) was employed to manage the coded texts and the analysis process indicates that data saturation was reached.

Ethical Considerations
The ethical review boards of UCSF and the Kenya Medical Research Institute approved the study protocol, data collection instruments, and consent forms. Approval was received from UCSF on May 22, 2018 and from Kenya Medical Research Institute on July 10, 2018. Verbal informed consent was obtained from all participants before study participation.

RESULTS
Social Franchise and SafeCare Training Benefits
Franchised private providers reported benefiting from trainings and mentorship on reproductive maternal and child health offered by the franchise networks. The overall objective of the training was to strengthen and improve knowledge of infection prevention and maternal and child health, and many providers felt they learned useful new skills through the trainings.

After joining AMUA, that’s when I went for training for family planning. Yeah, I was not competent on family planning before . . . After joining AMUA that is when I got introduced [to] those services and after being trained and I was given certificates. —R4PF07

Providers reported utilizing these acquired skills to support their clinic staff with continuous medical learning in their clinics. In some cases, providers worked with the quality assurance technical teams offered by the franchise to assess areas where improvement was needed and provided specific mentorship.

After the training we conducted internal CME [continuing medical education] where you provide feedback on what you learned on the daily management of clients. In some cases, we also invited one of our quality
assurance officers for a return on CME and we did resuscitation. We did CME training on resuscitations now for an adult, so it’s actually benefiting. —R4PF08

The majority of the franchised private providers reported that quality improvement training from the SafeCare program also had a positive impact on how they offered maternal and child health and family planning services in their clinics. Although trainings offered by the franchise networks focused more on skills-building, these were complemented by training that providers received through SafeCare, which helped them to update their standards for clinic operations. This training ranged from how to stock drugs properly to reduce waste, which types of equipment to purchase to enhance treatment safety and efficacy, and how to implement infection prevention practices.

Commodity Supply and Franchise Benefits

After joining a franchise, participating providers reported benefitting from steady family planning commodity supply, as well as access to equipment. Providers appreciated that the franchisors offered discounted prices on commodities and equipment that they felt improved service delivery, but otherwise would have been difficult to procure on the open market.

Like the benefits, you see when [we] were joining we had nothing [in] terms of commodities. AMUA supplied us with this coach, beds, gloves, and autoclaves. These commodities are hard to get in the market, they are expensive, but for us, we got them supplied at a lower price. —R4PF07

Private providers also said they routinely received family planning commodities from the franchisors to facilitate family planning service provision. The providers felt that this steady supply of commodities made it possible for them to offer quality care services, including offering additional services, such as medical male circumcision.

We benefited from family planning consumables such as gloves, jadels, femiplan pills. Those are now the commodities that were supplied. Even [voluntary medical male circumcision] commodities were also supplied too. In fact, our first starter pack for VMMC came from TUNZA; we were given three of them to start voluntary medical male circumcision. —R4PF29

Conversely, some providers complained that the Tunza franchise had a monopoly on a particular brand of contraceptive pill. Several clinics had special arrangements with suppliers from Tunza to stock this brand of pill. However, under this arrangement and despite demand from clients, it became difficult for providers to stock competing brands. This resulted in fewer family planning options for women visiting the clinic.

Experiences with Social Health Insurance

Beginning in 2016, franchised providers were given assistance to register and become accredited with the NHIF. Although both franchised and non-franchised providers reported many challenges with the NHIF accreditation process on their own, virtually all franchised providers who received this intervention found it beneficial and some reported that they would not have seen the accreditation process through without assistance from AHME.

In terms of improving quality, providers suggested that both preparing for accreditation and becoming accredited encouraged quality improvement in their facility. Not only did practice inspections conducted by the franchise representatives help providers to ensure quality compliance just before beginning the accreditation process, but a number of providers said that participating in SafeCare over the longer term enabled them to raise overall quality standards in their clinic even before the pre-accreditation site inspection. Technical assistance provided by SafeCare helped providers better adhere to Ministry of Health guidelines and follow through on quality improvement action plans, which in turn made the accreditation process easier.

You know when you improve quality, then your facility will receive NHIF representative accrediting your services, which is good. Those are some of the areas that we see assistance from the [SafeCare] teams. And then we feel they have really supported us by introducing action plans on areas that need improvement with a specific time frame. Yeah, you feel supported and of course, you agree to implement what has been agreed by the quality improvement team and that is very important. Those are some of the benefits. —RFPF25

Further, many franchised providers noted that NHIF accreditation was complex and difficult to navigate due to a lack of transparency around the process itself, challenges communicating with the local NHIF office, and corruption. The franchisors played a key role in helping providers to complete a process they might otherwise have abandoned.

NHIF assistance was very positive and very good, because AMUA is the ones who have put us here. Without them it would have been challenging and probably we would not have been accredited if it was not their support. They gave us a lot of support. —R4PF10
Notably, some nonfranchised facilities faced similar challenges and ultimately stopped pursuing accreditation as a result.

I have never known the reason [why my facility was not accredited]. I kept hoping, I kept ringing, but nobody gave me the reason. Then again, I said, I am an entrepreneur, they are not the people who brought me up to where I am, and I just stopped bothering them and moved on. —R4PN10

Once accredited, providers noted that participating in NHIF had allowed them to further improve quality by expanding their service offerings, which benefited patients in addition to benefiting the clinic as a business.

Even services the ones that we offer I see us improving as we go on, because before I came here there was nothing like CT scan. Now we have it because NHIF covers it. Even if it is things to do with the examinations, I told you things to do with the pictures like X-RAY. You see through the NHIF we can also advance business-wise because we have our machines. —R4PN19

**Strengthening Business Support**

The business support intervention assisted providers with quality by offering routine mentorship on financial management, record keeping, and drug and stock management. For example, some providers bought computer software applications that tracked drug expiry dates and assisted their clinic staff with auditing stock inventory, which helped these providers maintain a steady supply of drugs.

We have improved . . . because we are now keeping record of our drugs so . . . at the end of the day we have to know . . . how much drugs we have spent and what is remaining in the stock so that we can place an order immediately at least it has improved. —R4PF40

Although most providers spoke highly of SafeCare, they also encountered several challenges implementing quality improvement programs. Many private providers reported that it was difficult for them to expand or make structural changes to their clinic space to comply with SafeCare and that these changes required finances that were not easily available. Indeed, many providers reported that implementing SafeCare was expensive.

You know, we had challenges especially when you were on rented a room, you try to maintain some standard. I mean, for example I may like to put the tiles, but when you think about the cost of tiles and rent. You take a break and ask the landlord who will never do it. Because it will less on his rental income when he does the tiles. In the end, you just do it. It is a challenge to maintain SafeCare standards. —RFPF23

In some cases, providers were able to solve their financial capital challenges through other components of the AHME interventions. For example, NHIF accreditation enabled some providers to bring in more money by expanding their service offerings. For other providers, the acquisition of loans facilitated by the MCF allowed them to upgrade their clinics by purchasing new medical equipment that aided in offering comprehensive reproductive health services.

The major loan I took was through MCF and it was primarily for major upgrades of the facility. This facility is not what it was in 2013, everything has changed as I told you. So, most of the major things that you see here were done with the MCF funds. —R4PF24

However, while some providers appreciated the MCF loans and used them to improve their facilities, others had their own reservations regarding the interest rates and the loan repayment period. It is worth noting, though, that Kenyan interest rates were standardized in late 2016, which meant that rates negotiated through MCF that may once have been competitive were no longer more attractive than a standard bank loan. In addition, some providers were very skeptical of how they would transition through different stages of SafeCare while paying off bank loans they felt were at a high interest rate. These providers therefore had to weigh the benefits of one set of quality improvement plans against another.

**DISCUSSION**

This qualitative evaluation assessment provides rich data on private providers’ experiences with a set of quality improvement interventions under the umbrella of a program designed to improve access to quality health care services within private health facilities in Kenya. Quality of care interventions have shown the potential to improve reproductive and maternal-child health interventions in Kenya.38 However, quality of care interventions should be viewed in light of private providers’ needs to recoup their return on investments, cost, structural barriers, and limited access to loans for purposes of continued quality of care sustainability.24 Franchised providers felt that mentorship and capacity building on quality improvement offered through several components of the AHME intervention package improved their knowledge and
enabled them to provide a wider variety of higher quality services. Similarly, providers associated with NHIF accreditation with quality and several providers suggested that they would not have become accredited without support from AHME. However, providers often noted that quality improvement was costly both to implement and to maintain. Although some aspects of the AHME intervention package (e.g., MCF loans, increased client flow due to franchising) mitigated this challenge for some providers, cost was still a common concern among participating providers.

Our evaluation findings are similar to a number of other studies examining quality improvement in LMICs. Regarding NHIF accreditation, we found that private providers complained of an unclear application process that was lengthy and complex. In some cases, the complexity of this process deterred providers from applying for accreditation at all. This is consistent with findings from previous rounds of data collection for the AHME qualitative evaluation. Since NHIF accreditation is a means of quality assurance, deterring private providers from participating in this process through complexity and lack of transparency has implications for quality control among the private health sector in Kenya. Our findings showed that adhering to SafeCare requirements posed a significant challenge to private health care providers operating with little access to capital, particularly those located in low-income communities. Alkhenizan et al. also found that financial burdens imposed on health care facilities created barriers to quality improvement in LMICs. As Agha also found, our evidence suggests that offering loans to private providers is one way to decrease this financial burden for private providers in small facilities and increase quality of care. However, the loan themselves sometimes increased the providers’ financial burden, and providers often didn’t want to take loans because they were worried about interest rates and repayment.

Although public health providers have core expenses, such as rent and salaries, covered directly by government, private providers must rely only on their income to pay for all facility expenses. Private providers in small and medium-sized facilities in LMICs like Kenya often operate on very low budgets with little money left over. As shown by our findings, these tight budgets can affect clinical quality. Interventions like the AHME business support intervention are important for helping these providers learn to manage their finances and maintain a sustainable business that can afford to maintain and improve quality. In addition, some providers noted that joining NHIF allowed them to make improvements to their facilities, which in turn helped them generate more income to be put back into clinic upkeep. These findings combined with those around loans for private providers suggest a need for further research around the return on investment offered by similar quality improvement programs.

**Limitations**

These results should be viewed in light of the study’s limitations. Social desirability bias could have influenced the responses for both franchised and nonfranchised private providers, although it is difficult to predict the effect that social desirability bias would have had on these results. We made attempts to mitigate the potential effects of social desirability through the use of trained field interviewers and by emphasizing to participants that the interviewers were not representatives of any of the AHME partner organizations. Further, we note that the findings presented above would be richer if triangulated with other data sources, such as quantitative data on the extent to which provider quality actually improved through the AHME interventions. However, an external quantitative evaluation meant to complement the qualitative results presented here was delayed such that the qualitative and quantitative teams were not able to cross-reference their findings. Despite these limitations, we feel that this article provides novel insights on experiences of franchised private providers reporting their experiences with quality improvement interventions.

**CONCLUSION**

Several studies have shown evidence that social franchising models have worked to improve the overall health outcomes of their communities through the quality of care interventions. Our findings suggest that engaging private providers in efforts to improve quality of care in private clinics through a package of interventions that extend beyond the typical social franchising model is achievable. Further, this model may be preferable to traditional social franchising where possible because it offers more customizability to meet the needs of private providers across a range of facility sizes and income levels. However, because weighing the benefits of quality improvement against the costs of implementing a comprehensive quality improvement program remained a critical concern for private providers, we recommend further
research on the return on investment that such quality improvement interventions can offer to private providers in small and medium-sized facilities in LMICs.

Acknowledgments: We would like to thank Avery Seefeld and Ginger Galub for their research and administrative support on this study. We acknowledge all the participating franchise Tunza and Amua facilities across the county in Kenya and the overwhelming support of study participants.

Funding: This study was jointly funded by the Bill and Melinda Gates Foundation and the United Kingdom Department for International Development. The funders had no role in the study design, implementation, or decision to publish.

Conflicting interests: None declared.

REFERENCES

sustainable-development-goals/
26. Campbell AM, Montagu D, Prata N, Orero S, Walsh J. Kenya: Reaching the Poor Through the Private Sector - A Network Model for


Private Providers’ Experiences Implementing Quality of Care Programs in Kenya

Peer Reviewed

Received: January 19, 2020; Accepted: July 7, 2020; First published online: August 17, 2020


© Syengo and Suchman. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are properly cited. To view a copy of the license, visit http://creativecommons.org/licenses/by/4.0/. When linking to this article, please use the following permanent link: https://doi.org/10.9745/GHSP-D-20-00034
Determinants of Facility-Level Use of Electronic Immunization Registries in Tanzania and Zambia: An Observational Analysis

Emily Carnahan, Ellen Ferriss, Emily Beylerian, François Dien Mwansa, Ngwegwe Bulula, Dafrossa Lyimo, Anna Kalbarczyk, Alain B. Labrique, Laurie Werner, Jessica C. Shearer

ABSTRACT

Background: As more countries transition from paper-based to electronic immunization registries (EIRs) to collect and track individual immunization data, guidance is needed for successful adoption and use of these systems. Little research is available on the determinants of EIR use soon after introduction. This observational study assesses the determinants of facility health care workers’ use of new EIRs in Tanzania and Zambia, implemented during 2016 to 2018.

Methods: We used EIR data entered between 2016 and 2018 from 3 regions in Tanzania and 1 province in Zambia to measure weekly EIR system use for a total of 50,639 facility-weeks. We joined secondary data on facility characteristics and applied the Performance of Routine Information System Management framework to categorize characteristics as organizational, technical, or behavioral. We used a generalized estimating equations logistic regression model to assess facility characteristics as potential determinants of system use.

Results: In both countries, the estimated odds of weekly EIR use declined weekly after EIR introduction. In Tanzania, health centers and hospitals had increased odds of system use compared to dispensaries. For each additional health care worker trained in a facility during the EIR introduction, the estimated odds of weekly EIR use increased. Tanzanian facilities that had transitioned entirely to paperless reporting had higher odds of sustained use compared to those maintaining parallel electronic and paper-based reporting systems. In Zambia, distance from the district health office was significantly associated with decreasing odds of system use. There were significant differences in EIR use by district in both countries.

Discussion: The results highlight the importance of organizational and behavioral factors in explaining sustained EIR use. As EIRs are introduced in new settings, we recommend indicators of engagement and use be built directly into the system for routine monitoring and implementing changes as needed.

INTRODUCTION

Each year, vaccines prevent an estimated 2 to 3 million child deaths.1 Over the last 3 decades, vaccination coverage levels have significantly increased worldwide, but progress has stalled in recent years. In 2018, 19.4 million children did not receive the 3 recommended doses of...
diptheria, tetanus, and pertussis vaccine in their first year of life. In low- and middle-income countries (LMICs), efforts to strengthen vaccination coverage and equity are often impeded by inaccurate or incomplete data on routine childhood immunizations. As many country health systems are integrating new digital tools, there is an opportunity to use digital tools to improve immunization programs’ pursuit to close the vaccination gap.

The electronic immunization registry (EIR) is one such digital tool to support immunization program performance. EIRs are routine systems to capture, store, access, and share individual, longitudinal health information in digitized records. They focus on collecting individual immunization records but can also collect other health and demographic data in each individual’s digital record and can be linked with vaccine stock management, human resources management, or other health management information systems. There is promising—although limited—evidence that digital health tools can improve vaccination adherence, uptake, and the efficiency of immunization programs. In high-income countries, where they are routinely used, EIRs have been shown to support tracking individual vaccine eligibility and delivery, facilitate vaccine management and accountability, and inform assessments of vaccination coverage, missed opportunities, and disparities, among other benefits. Thus, EIRs have been proposed as a solution to improve data quality, facilitate reporting, and promote data use in LMICs, ultimately providing the opportunity to strengthen vaccination services.

Traditionally, most LMICs have relied on paper-based data collection at the facility level to capture immunization data. As country demand to integrate digital solutions to improve health outcomes is increasing, there are also more examples of LMICs developing, implementing, and scaling EIRs across many countries in Latin America, Africa, and Asia. The Bill and Melinda Gates Foundation, Gavi, and other funders have invested substantially in EIRs in LMICs to improve immunization data quality and use to increase vaccine coverage and equity. However, few EIR implementations in LMICs have been rigorously evaluated. And despite improved political, financial, and technical support to introduce digital health strategies, many of these interventions continue to face barriers to adoption and consistent use that limit their potential impact.

As more countries transition from traditional systems using paper-based data collection tools to EIRs, guidance is needed to ensure the ongoing use of these new systems. This study uses an implementation science framework to test hypotheses of the drivers of EIR use. Understanding the drivers can inform improvements to the system design and/or implementation strategies for more effective implementation and sustained use. This article seeks to identify determinants of facility-level differences in system use among facilities in Tanzania and Zambia where EIRs have been implemented.

| THE EIR INTERVENTION AS PART OF TANZANIA AND ZAMBIA’S EHEALTH LANDSCAPE |

Tanzania and Zambia are the focus of this study based on their role as demonstration countries in the BID Initiative. Funded by the Bill and Melinda Gates Foundation, the BID Initiative is grounded in the belief that better data plus better decisions will lead to better health outcomes.

In Tanzania and Zambia, facility health care workers (HCWs) use paper-based tools to capture information about vaccine delivery and stock. Each facility manually aggregates the data into a monthly report that is submitted to the district; from there, data are entered into electronic systems and aggregated to the district, regional, and national levels. Beginning in 2013, the Tanzanian and Zambian Ministries of Health (MOHs), in partnership with PATH, identified challenges related to immunization data quality and use and then iteratively developed solutions to address them.

The most pressing immunization data quality and use challenges (categorized by the World Health Organization Classification of Digital Health Interventions health system challenges) included the following:

- Incomplete or untimely data (1.3)
- Inaccurate or uncertain population denominators to inform coverage calculations (1.1)
- Lack of unique identifiers for infants (1.7)
- Difficulty identifying children who do not start immunization or who drop out (defaulter tracing) (1.5, 8.6)
- Poor data visibility at the facility level (1.5)
- Complex data collection tools
- Insufficient data on supply chains and logistics management (1.3)
- Inadequate capacity for data management and use (1.6)
The BID Initiative intervention strategy to address these challenges included both technological and change management components to foster an environment conducive to data use for decision making.21 The package of solutions included the development of a standards-based EIR with automated, simplified reports; web-based dashboards; and supply chain system tools,12 and the introduction of data use mentors, peer networking communication forums between health workers, and data use guides to build capacity and motivation for data use.22 Interventions were designed according to the principles of user-centered design to address each country’s identified challenges.23,24 For example, in both countries, the MOHs codesigned requirements for the EIRs with the project team, and user advisory groups (comprised of HCWs from facility, district, and regional levels) tested iterations of the EIRs to provide feedback on the functionality and user interface.12

The BID team tested and refined interventions in pilot facilities in each country before scaling up a package of interventions to other facilities. The interventions were introduced to HCWs through on-the-job training with staff from higher levels of the health system (district, regional, national) engaged to provide a supportive environment through championing data use practices, mentoring facility staff, and holding facilities accountable for their performance.25

DETERMINANTS OF EIR USE: A CONCEPTUAL FRAMEWORK

We sought to evaluate facility characteristics associated with EIR use following the rollout of interventions in the first regions of implementation—Arusha, Kilimanjaro, and Tanga regions, Tanzania, and Southern Province, Zambia—from 2016 to 2018. The successful adoption of EIRs requires many organizational, technical, and behavioral factors to come together, as conceptualized in the Performance of Routine Information System Management framework.26 Facility characteristics in this study were mapped to these categories of determinants. The hypothesized impact of these characteristics on EIR use are shown in Table 1 and described here.

Organizational Factors

Organizational factors included the level of supervisory and political support for the new system, availability of human and financial resources, and management support. These factors were manifested at the facility, district, or regional level through informal norms, values, and practices or through formal guidelines, standards, and policies. As EIRs were introduced, organizational policies often required HCWs to continue the traditional paper-based data entry in addition to entering data electronically; this was the case in Tanzania and Zambia where HCWs were expected to conduct parallel data entry in the EIR and official paper-based reporting system until officially switching to paperless reporting. We hypothesized that when a facility transitioned to paperless reporting (i.e., only using the EIR), system use would have increased. A study evaluating SmartCare EIR use in Zambia observed parallel data entry requirements undermined SmartCare EIR adoption by clinic staff.27 We hypothesized based on implementation experience that HCWs in facilities responsible for a larger population would have had a busier daily client load, therefore less time for data entry. A study of the Mobile Technology for Health program in Ghana found that higher volume health centers and hospitals were less likely than community-based facilities to register and upload individual-level health information to the mobile platform, in part due to overburdened HCWs.28 However, we also hypothesized the reverse could be true: that HCWs in facilities with a larger patient population may have seen more value in using the EIR to manage their patients and would have been more likely to use the system. We also hypothesized relationships between EIR use and facility type, ownership, distance to the district health office (DHO) (as a proxy for supportive supervision), training strategy, and number of immunization sessions offered per week (Table 1).

Technical Factors

Technical factors, such as user-interface design and offline functionality, were likely to affect the user’s experience with an EIR system as well as the system’s feasibility and acceptability. Considering that the system itself was the same across facilities, we hypothesized that facilities with a consistent connection to internet and electricity would have been more likely to use the system. Other studies have shown that EIR adoption in 2 districts in Uganda was impeded by blackout days (no electricity or internet connectivity)29 and that power outages (“load-shedding” or brown-outs) were the primary challenge to using the SmartCare electronic health record system for immunization data in Zambia.27 Technical factors could have affected system performance directly or mediated through behavioral factors. For example, in Zambia, clinic
Behavioral Factors

Behavioral factors, such as HCWs’ capability and motivation to use the new system, required careful attention during system design, implementation, and beyond. We hypothesized that HCWs who received adequate training and more or higher quality ongoing supportive supervision would have been more likely to use the EIR. A review of an EIR introduction in Uganda highlighted the importance of onsite technical support and on-the-job training.\(^2^9\) We hypothesized that if more HCWs in a facility had been trained to use the EIR, the facility would have been more likely to have sustained use of the EIR over time. We also hypothesized that HCW motivation to use the EIR would possibly wane over time, so as the length of time since EIR introduction increased, facilities would have been less likely to use the EIR. A recent systematic review of data use interventions found that HCWs were not motivated to adopt or use new digital interventions when they replaced a status quo that was perceived to work adequately.\(^9\) Others have suggested the perceived threat of increased data transparency (and thereby potential scrutiny) may limit enthusiasm for digital systems.\(^3^0\) Although other behavioral factors (e.g., HCWs’ attitudes, skills, and motivation) are important determinants, this study was limited to including behavioral factors that could be measured using existing secondary data sources.

Additional Factors

Other factors, such as national leadership, governance, and policy, are important aspects of the enabling environment for any digital tool\(^3^0,^3^1\); however, they were not included in this study.

---

**TABLE 1. Hypotheses on Impact of Facility Characteristics on EIR Use Aligned to PRISM Framework**

<table>
<thead>
<tr>
<th>PRISM Framework Determinant</th>
<th>Variable</th>
<th>Hypotheses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organizational Paperless reporting</td>
<td>If a facility transitions to paperless reporting (only using the EIR as the official system), it will be more likely to use the EIR.</td>
<td></td>
</tr>
<tr>
<td>Facility volume</td>
<td>If a facility has a larger patient population, HCWs may have a busier daily patient load, therefore less time for data entry and will be less likely to use the EIR. OR, if a facility has a larger patient population, HCWs may see more value in using the EIR to manage their patient population and will be more likely to use the EIR.</td>
<td></td>
</tr>
<tr>
<td>Facility type</td>
<td>If a facility is a hospital or health center, it may have more resources (e.g., equipment, skilled HCWs) compared to a dispensary and may be more likely to use the EIR.</td>
<td></td>
</tr>
<tr>
<td>Ownership type</td>
<td>If a facility is public, HCWs may feel greater ownership of the decision to adopt the EIR and/or feel more accountable to use the EIR than in private facilities and thus may be more likely to use the EIR.</td>
<td></td>
</tr>
<tr>
<td>Distance to district health office</td>
<td>If a facility is located closer to the district health office, it will be more likely to receive in-person support from district health officials.</td>
<td></td>
</tr>
<tr>
<td>Training strategy (Tanzania only)</td>
<td>If a facility received the second training strategy (i.e., district staff provided additional support and training), it will be more likely to use the EIR than facilities who received the first training strategy, which relied on BID project staff.</td>
<td></td>
</tr>
<tr>
<td>Number of immunization sessions per week</td>
<td>If a facility provides more immunization sessions per week, they will be more likely to enter data into the EIR each week.</td>
<td></td>
</tr>
<tr>
<td>Technical Primary power source</td>
<td>If a facility has a consistent electricity connection, it will be more likely to use the EIR.</td>
<td></td>
</tr>
<tr>
<td>Internet connectivity</td>
<td>If a facility has a consistent internet connection, it will be more likely to use the EIR.</td>
<td></td>
</tr>
<tr>
<td>Behavioral Number of HCWs trained per facility</td>
<td>If a facility has more HCWs trained, it will be more likely to use the EIR.</td>
<td></td>
</tr>
<tr>
<td>Weeks since EIR introduction</td>
<td>As the length of time since EIR introduction increases, facilities will be less likely to use the EIR.</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: EIR, electronic immunization registry; HCW, health care worker; PRISM, Performance of Routine Information System Management.
due to the focus on determinants at the facility level.

**METHODS**

**Study Setting**

The study data are from electronic vaccine records in 3 regions of Tanzania: Arusha, Kilimanjaro, and Tanga, and in Zambia’s Southern Province, where the BID Initiative interventions were initially implemented. The regions in Tanzania were chosen by the government as the first implementation regions due to their nomadic communities and porous borders with Kenya, which posed challenges to estimating population denominators for monitoring immunization coverage. Southern Province in Zambia had similar challenges due to sharing a border with Zimbabwe in the southwest and Botswana in the southeast. The regions in Tanzania and Zambia were also chosen for their mix of urban, peri-urban, and rural communities. Moreover, Southern Province in Zambia was selected as the pilot province due to underperformance of the immunization program and strong commitment from the district and provincial leaders.

The 3 regions in Tanzania collectively have a population of 5.38 million, cover a land mass of over 77,500 square kilometers, and support 847 health facilities that provide immunization services. Southern Province in Zambia has a population of 1.96 million, covers a land mass of 85,283 square kilometers, and supports 298 health facilities that provide immunization services.

**Data**

We used data routinely collected through the EIRs in Tanzania and Zambia. Tanzania first introduced a system called the Tanzania Immunization Information System in Arusha region in 2016–2017, which was later replaced with an improved system called the Tanzania Immunization Registry (TImR) built upon the OpenIZ platform (now known as SanteDB). Rollout was phased across 26 districts beginning in the Tanga, Arusha, and Kilimanjaro regions. Zambia introduced a system called the Zambia Electronic Immunization Registry (ZEIR) that was built upon the OpenSRP platform beginning in Southern Province with a phased rollout across the 13 districts. TImR and ZEIR are both open-source software developed for use on tablets with online and offline functionality. In both countries, paper-based reporting remained the official system for capturing childhood immunization status until facilities fully transitioned to paperless reporting.

All data collected via Tanzania and Zambia’s routine immunization programs were released for analysis with permission from the Governments of Tanzania and Zambia. Institutional review board approval was obtained from both countries and a non-human subjects research determination was received from the PATH institutional review board. The analysis team received anonymized digital data. Data were processed and analyzed using Tableau, Alteryx, R, and STATA.

Data from the 3 regions in Tanzania were extracted from TImR on June 11, 2018, and data from Southern Province in Zambia were extracted from ZEIR on October 3, 2018. We excluded data for the final (partial) month and penultimate month to avoid completeness issues as a result of facilities working in the offline mode that had not yet synced their data for the most recent months. We included all facilities that had ever entered data into the EIR, thereby excluding facilities that did not provide immunizations. Table 2 provides descriptive information about the EIR datasets. Each data record was a single service delivered to an individual, and an individual could have had multiple services delivered during a single visit. Services included child weights captured, child or adult vaccines delivered, and nonvaccine interventions delivered (e.g., vitamin A, mebendazole, or insecticide-treated nets).

Datasets were cleaned for analysis to exclude historical records (back-entered data) by only considering data entries with a date after the EIR had been introduced in each facility. This eliminated the effect of large batches of backlogged records submitted by individual facilities shortly after the EIR deployment.

Secondary data on facility characteristics were collected by the BID Initiative implementation team during EIR rollout. Facility types and ownership were categorized based on government definitions. Distance from the DHO was calculated using facility GPS coordinates. Data on facility volume were extracted from government health management information systems.

**Measures of System Use**

We used the dates of data entry into the EIR as a proxy for system use. We measured system use on a weekly basis for each facility based on whether there was any EIR activity at that facility, as measured by whether data records capturing any events (e.g., vaccine dose delivered, child weight
recorded) were entered into the EIR for the given week. Our outcome variable for system use was binary, with 1 indicating that at least 1 data record had been entered into the EIR for that facility-week.

**Analysis**

We used generalized estimating equation logistic regression models (xtgee in STATA with AR[1] correlation structure) to obtain marginal estimates of the odds ratios associated with facility characteristics and to account for within-facility correlation arising from longitudinal panel data with multiple observations (weeks) for each facility (the panel identifier). Covariates were chosen based on our conceptual framework, existing literature, and the authors’ firsthand experience participating in the BID Initiative implementation.

Organizational covariates included the facility type, ownership (public or private), volume of service delivery (number of vaccines administered monthly in Tanzania and annual number in attendance at child health clinics in Zambia), number of immunization sessions provided per week (Zambia only), and an indicator for whether the facility had transitioned to paperless reporting (Tanzania only). Distance from the DHO was included as a proxy for support from district health officials, as we hypothesized that they were more likely to provide in-person support to facilities that are physically closer. The training strategy provided in Tanzania was also tested as an organizational covariate, as midway through implementation in the first region, the strategy shifted from on-the-job training provided by BID project staff to leveraging district staff to provide additional support and training. Technical covariates included the facility’s primary power source and internet connectivity. Behavioral covariates included the number of HCWs initially trained during implementation rollout and the total number of weeks since the EIR was introduced at the facility. The data sources for covariates included facility characteristics captured in the EIRs, BID Initiative program data, government administrative data, and historical vaccine delivery data from paper-based sources.

Results were modelled separately for each country using complete case analysis. Covariates were first tested for a bivariate significant relationship with the outcome variable, and those with statistical significance were included in the final regression model. The final model for Tanzania included time since EIR introduction, facility type, number of HCWs trained on the EIR, and an indicator for when the facility transitioned to paperless reporting as covariates. The final model for Zambia included time since EIR introduction, distance from the DHO, and the number of immunization sessions provided per week. The final models for both countries additionally included the district where the facility was located. We hypothesized system use would differ by district due to a host of organizational, technical, and behavioral factors, both captured and not captured in the final models. Districts were included to control for confounding between the previously mentioned covariates and EIR use resulting from

---

**TABLE 2. Description of the Datasets Extracted From the Tanzania Immunization Registry and Zambia Electronic Immunization Registry**

<table>
<thead>
<tr>
<th></th>
<th>Arusha Region</th>
<th>Kilimanjaro Region</th>
<th>Tanga Region</th>
<th>Southern Province</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of districts</td>
<td>6</td>
<td>6</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>Number of facilities</td>
<td>283</td>
<td>292</td>
<td>330</td>
<td>551 (static and outreach sites)</td>
</tr>
<tr>
<td>Number of unique individuals</td>
<td>137,130</td>
<td>35,084</td>
<td>89,740</td>
<td>96,617</td>
</tr>
<tr>
<td>Number of records</td>
<td>1,606,776</td>
<td>206,871</td>
<td>671,562</td>
<td>1,323,264</td>
</tr>
<tr>
<td>Date range of EIR records (including back-entered data)</td>
<td>January 2015 – April 2018</td>
<td>January 2015 – April 2018</td>
<td>January 2015 – April 2018</td>
<td>January 2015 – August 2018</td>
</tr>
<tr>
<td>Date range of EIR introduction</td>
<td>June 2016 – March 2017</td>
<td>December 2017 – February 2018</td>
<td>July 2017 – August 2017</td>
<td>July 2017 – March 2018</td>
</tr>
</tbody>
</table>

Abbreviation: EIR, electronic immunization registry.
unmeasured facility characteristics imparted by the district in which the facility was located. The district that was expected to demonstrate the highest performance was selected as the comparator group in each model. In Tanzania, district-level units included city councils (CC), district councils (DC), municipal councils (MC), or town councils (TC). For the Tanzania model, Arusha CC in Arusha region was selected, as it was the pilot implementation district, thus had received additional support, and includes the capital and largest city in Arusha region. For the Zambia model, Choma district was selected as it is the capital district for Southern Province.

RESULTS

Activity

In Tanzania, the package of interventions was implemented in 285 facilities in Arusha, 327 facilities in Tanga, and 312 facilities in Kilimanjaro region. Figure 1 shows the cumulative number of facilities using the EIR based on staged introduction, and of those facilities, the percentage that used the system in each week. In Arusha region, the number of facilities where the EIR had been introduced gradually increased from June 2016 to March 2017 as this was the pilot region where interventions were being iteratively adapted and gradually introduced to new districts, ultimately resulting in 278 unique facilities using the system. There was a much more rapid scale-up in Kilimanjaro and Tanga regions, where interventions were introduced across all districts within 2–3 months. In Tanga region, 325 facilities had ever entered data into TImR, and in Kilimanjaro, 285 facilities, similarly, had entered any data. Across all regions, the number of active facilities plateaued and then declined over time. We calculated the number of active weeks as a percent of total weeks since EIR introduction by facility: Figure 2 shows the facility average by district in Tanzania. The district average ranges from 39% in Kilindi (meaning on average, facilities in Kilindi used the EIR for 39% of the total weeks since introduction) to 86% in Tanga district within the Tanga region.

In Zambia, the package of interventions was introduced in 298 facilities across 13 districts in Southern Province beginning in November 2016 in Livingstone District. ZEIR was first introduced in Livingstone and Kazungula Districts in July 2017 and was scaled up to the other districts through March 2018. Figure 3 shows the number of active facilities over time as the EIR was introduced. Like Tanzania, there was a gradual increase in use as the EIR was introduced to new districts, but then system use began to decline over time. The steady decline in EIR use was similar across districts (results not shown). In August 2018, fewer than 5% of facilities where the EIR had been introduced in Southern Province had entered data.
Figure 2. Facility Average Percentage of Active Weeks of EIR Use by District, Tanzania, 2016–2018

Figure 4 shows the facility average percentage of weeks active in the EIR by district. The district average ranges from 20% in Siavonga to 54% in Choma.

Facility Determinants of Weekly Activity
The regression models (1 per country) explain the relative contribution of the facility- and district-level determinants associated with weekly EIR use. Table 3 describes the independent variables that were tested for inclusion in the models, and Tables 4 and 5 show the model results for each country.

Tanzania
In Tanzania, 2 organizational determinants, facility type and whether the facility had transitioned fully to paperless reporting were significant predictors of EIR use (Table 4). Compared to dispensaries, health centers were 61% more likely (odds ratio [OR]=1.61; 95% confidence interval [CI]=1.08, 2.42) to use the system, while hospital odds of use were 3.83 (95% CI=2.14, 6.85) times greater. Facilities that had transitioned to completely paperless reporting had odds of weekly EIR use that were 2.76 (95% CI=1.54, 4.94) times as large as facilities using parallel EIR and paper reporting systems. The log-transformed median number of doses delivered per month, a measure of facility volume, was found to be significantly associated with EIR use in the bivariate model (OR=1.76; 95% CI=1.57, 1.98); however, this was excluded from the final model due to collinearity with facility type and number of health workers trained. Ownership type, distance to the DHO, and the training strategy received, were not found to be significantly associated with EIR use in bivariate analyses, thus were excluded.

Technical determinants of EIR use were not included in the final model. In bivariate analysis, facilities with no primary power source were estimated to have significantly lower weekly EIR use
compared to facilities that were connected to the electric grid (OR=0.35; 95% CI=0.15, 0.81). However, primary power source was excluded from the final model as data were not available for Tanga region. Internet connectivity was not significantly associated with system use in bivariate analysis, thus was also excluded.

Behavioral determinants significantly associated with EIR use included the number of HCWs trained and weeks since EIR introduction. For each additional HCW that was trained during the EIR introduction, estimated odds of weekly EIR use were 1.39 (95% CI=1.22, 1.58) times greater. For each additional week from EIR introduction, the odds of use were 1.9% lower (95% CI=1.5%, 2.3%) (OR=0.98; 95% CI=0.98, 0.99).

Facilities in most districts did not have significantly different odds of using the EIR compared to facilities in Arusha CC. The exceptions were facilities in Tanga CC and Karatu DC, which were estimated to have significantly higher odds of use compared to facilities in Arusha CC, and facilities in Same DC, Muheza DC, Kilindi DC, and Moshi DC, which were estimated to have significantly lower odds of use compared to facilities in Arusha CC.

Zambia
In Zambia, distance from the DHO was the only organizational determinant significantly associated with EIR use (Table 5). Facilities in the second distance quartile were estimated to have 0.46 times (95% CI=0.24, 0.86) the odds of weekly use compared to facilities in the first quartile. Facilities in the third and fourth quartiles were estimated to have 0.41 (95% CI=0.21, 0.79) and 0.32 times (95% CI=0.17, 0.63) the odds of weekly use compared to facilities in the first quartile for distance to the DHO. Urban/rural status was significantly associated with EIR use, however was collinear with distance to the DHO, thus excluded from the final model. Odds of use were not significantly different between facilities offering vaccination days less than once a week and those offering vaccination days at least once a week. Facility volume (measured as 2017 attendance at child health clinics) and facility type were not found to be significantly associated with EIR use in bivariate analyses, thus were excluded from the model.

The only technical determinant assessed for Zambia, primary power source, was not found to be statistically significantly associated with EIR use, thus was omitted from the final model.

The only behavioral covariate significantly associated with EIR use was time since EIR introduction. The odds of EIR use decreased by 12.4% (95% CI=11.5, 13.4) per week from introduction.

Districts with significantly lower odds of estimated weekly EIR use included Gwembe, Kalomo, Namwala, Sinazongwe, Pembu, Monze, Chikankata, and Siavonga. Odds of use were not significantly different in Zimba, Kazungula, Mazabuka, or Livingstone districts compared to Choma.
DISCUSSION

The descriptive analyses based on facility EIR data showed declines in weekly EIR use post-introduction across all regions in Tanzania and Zambia. The statistical analyses joined the EIR data with secondary data on other facility characteristics to explain which facility determinants were associated with system use.

Organizational factors that were strongly associated with weekly EIR use were facility type and paperless reporting in Tanzania and distance to the DHO in Zambia. The results show higher odds of system use associated with hospitals and health centers compared to dispensaries in Tanzania. This may have been because larger facilities were more likely to have adequate HCWs dedicated to immunization of whom some could prioritize data entry, as opposed to dispensaries where a single HCW would likely be stretched across functional areas. HCWs in hospitals or health centers may have had greater technical skills, training, and capacity compared to HCWs in dispensaries so they may have been more adept at using the system. HCWs in larger facilities may have also perceived the EIR as more valuable to support their day-to-day work to manage and track a large client population—a task that would be more manageable with paper-based forms if the client population was smaller.

Facilities in Tanzania that had transitioned to paperless reporting were significantly more likely to use the EIR compared to facilities that were still responsible for parallel systems. BID staff in both countries observed that when HCWs had limited bandwidth, they prioritized data entry in the paper-based tools (the official reporting system) over using the EIR. Once a facility transitioned to using the EIR as their official
<table>
<thead>
<tr>
<th>Description</th>
<th>Arusha</th>
<th>Kilimanjaro</th>
<th>Tanga</th>
<th>All Regions</th>
<th>Southern Province</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of districts</td>
<td>6</td>
<td>6</td>
<td>8</td>
<td>20</td>
<td>13</td>
</tr>
<tr>
<td>Number of facilities</td>
<td>278</td>
<td>285</td>
<td>326</td>
<td>889</td>
<td>282</td>
</tr>
</tbody>
</table>

### Organizational

<table>
<thead>
<tr>
<th>Description</th>
<th>Tanzania</th>
<th>Zambia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper-based records</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Paperless records</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Number of monthly vaccine doses delivered, mean (SD)</td>
<td>341.3 (529.1)</td>
<td>206.2 (246.5)</td>
</tr>
<tr>
<td>Annual child health clinic attendance, mean (SD)</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

### Facility type

<table>
<thead>
<tr>
<th>Description</th>
<th>Arusha</th>
<th>Kilimanjaro</th>
<th>Tanga</th>
<th>All Regions</th>
<th>Southern Province</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dispensary</td>
<td>78.4%</td>
<td>79.3%</td>
<td>86.5%</td>
<td>81.7%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Health center</td>
<td>16.9%</td>
<td>15.8%</td>
<td>0.0%</td>
<td>10.3%</td>
<td>76.1%</td>
</tr>
<tr>
<td>Hospital</td>
<td>4.7%</td>
<td>4.9%</td>
<td>2.5%</td>
<td>3.9%</td>
<td>1.8%</td>
</tr>
<tr>
<td>Hospital affiliated center</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>4.4%</td>
</tr>
<tr>
<td>Missing</td>
<td>0.0%</td>
<td>0.0%</td>
<td>11.0%</td>
<td>4.0%</td>
<td>17.6%</td>
</tr>
</tbody>
</table>

### Ownership type

<table>
<thead>
<tr>
<th>Description</th>
<th>Arusha</th>
<th>Kilimanjaro</th>
<th>Tanga</th>
<th>All Regions</th>
<th>Southern Province</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private</td>
<td>32.4%</td>
<td>24.6%</td>
<td>12.0%</td>
<td>22.4%</td>
<td>–</td>
</tr>
<tr>
<td>Public</td>
<td>65.1%</td>
<td>70.5%</td>
<td>85.6%</td>
<td>74.4%</td>
<td>–</td>
</tr>
<tr>
<td>Missing</td>
<td>2.5%</td>
<td>4.9%</td>
<td>2.5%</td>
<td>3.3%</td>
<td>–</td>
</tr>
</tbody>
</table>

### Distance to DHO, km, mean (SD)

<table>
<thead>
<tr>
<th>Description</th>
<th>Arusha</th>
<th>Kilimanjaro</th>
<th>Tanga</th>
<th>All Regions</th>
<th>Southern Province</th>
</tr>
</thead>
<tbody>
<tr>
<td>37.2 (31.0)</td>
<td>61.8 (171.5)</td>
<td>23.4 (14.4)</td>
<td>35.9 (68.8)</td>
<td>46.7 (39.7)</td>
<td></td>
</tr>
</tbody>
</table>

### On-the-job training by BID Initiative staff

<table>
<thead>
<tr>
<th>Description</th>
<th>Arusha</th>
<th>Kilimanjaro</th>
<th>Tanga</th>
<th>All Regions</th>
<th>Southern Province</th>
</tr>
</thead>
<tbody>
<tr>
<td>71.6%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>22.4%</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

### Additional support and training by district staff

<table>
<thead>
<tr>
<th>Description</th>
<th>Arusha</th>
<th>Kilimanjaro</th>
<th>Tanga</th>
<th>All Regions</th>
<th>Southern Province</th>
</tr>
</thead>
<tbody>
<tr>
<td>28.4%</td>
<td>100.0%</td>
<td>100.0%</td>
<td>77.6%</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

### Number of immunization sessions per week

<table>
<thead>
<tr>
<th>Description</th>
<th>Arusha</th>
<th>Kilimanjaro</th>
<th>Tanga</th>
<th>All Regions</th>
<th>Southern Province</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 or more</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>77.4%</td>
</tr>
<tr>
<td>Less than 1</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>11.0%</td>
</tr>
<tr>
<td>Missing information</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>11.6%</td>
</tr>
</tbody>
</table>

### Technical

<table>
<thead>
<tr>
<th>Description</th>
<th>Arusha</th>
<th>Kilimanjaro</th>
<th>Tanga</th>
<th>All Regions</th>
<th>Southern Province</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary power source</td>
<td>Grid</td>
<td>36.7%</td>
<td>79.3%</td>
<td>0.0%</td>
<td>36.9%</td>
</tr>
<tr>
<td>Solar</td>
<td>31.3%</td>
<td>7.7%</td>
<td>0.0%</td>
<td>12.3%</td>
<td>1.8%</td>
</tr>
<tr>
<td>None</td>
<td>4.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>1.2%</td>
<td>51.1%</td>
</tr>
<tr>
<td>Missing</td>
<td>28.1%</td>
<td>13.0%</td>
<td>100%</td>
<td>49.6%</td>
<td>3.5%</td>
</tr>
</tbody>
</table>

### Internet connectivity

<table>
<thead>
<tr>
<th>Description</th>
<th>Arusha</th>
<th>Kilimanjaro</th>
<th>Tanga</th>
<th>All Regions</th>
<th>Southern Province</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>60.8%</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>No</td>
<td>6.8%</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Missing</td>
<td>32.4%</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

### Behavioral

<table>
<thead>
<tr>
<th>Description</th>
<th>Arusha</th>
<th>Kilimanjaro</th>
<th>Tanga</th>
<th>All Regions</th>
<th>Southern Province</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of HCWs trained per facility, mean (SD)</td>
<td>2.2 (0.9)</td>
<td>2.1 (1.2)</td>
<td>2.5 (1.1)</td>
<td>2.3 (1.1)</td>
<td>–</td>
</tr>
</tbody>
</table>

Abbreviations: DHO, district health office; EIR, electronic immunization registry; HCW, health care worker; SD standard deviation.
### TABLE 4. Results of Regression Model Predicting EIR Use for Facilities in Tanzania

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate (odds ratio)</th>
<th>Robust Standard Error</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Organizational</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paperless (compared to using parallel systems)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.72</td>
<td>0.83</td>
<td>.001</td>
</tr>
<tr>
<td>Facility Type (compared to dispensary)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health center&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.61</td>
<td>0.33</td>
<td>.02</td>
</tr>
<tr>
<td>Hospital&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.82</td>
<td>1.13</td>
<td>&lt;.001</td>
</tr>
<tr>
<td><strong>Behavioral</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of HCWs trained&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.35</td>
<td>0.09</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weeks since EIR introduction&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.98</td>
<td>&lt;0.01</td>
<td>&lt;.001</td>
</tr>
<tr>
<td><strong>District (Region)&lt;sup&gt;b&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tanga CC (Tanga)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.89</td>
<td>1.07</td>
<td>.004</td>
</tr>
<tr>
<td>Karatu DC (Arusha)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.45</td>
<td>0.81</td>
<td>.007</td>
</tr>
<tr>
<td>Mkinyinga DC (Tanga)</td>
<td>1.83</td>
<td>0.64</td>
<td>.09</td>
</tr>
<tr>
<td>Pangani DC (Tanga)</td>
<td>1.56</td>
<td>0.54</td>
<td>.20</td>
</tr>
<tr>
<td>Longido DC (Arusha)</td>
<td>1.53</td>
<td>0.65</td>
<td>.32</td>
</tr>
<tr>
<td>Ngorongoro DC (Arusha)</td>
<td>1.53</td>
<td>0.56</td>
<td>.24</td>
</tr>
<tr>
<td>Handeni TC (Tanga)</td>
<td>1.32</td>
<td>0.70</td>
<td>.60</td>
</tr>
<tr>
<td>Korogwe TC (Tanga)</td>
<td>1.19</td>
<td>0.62</td>
<td>.74</td>
</tr>
<tr>
<td>Siha DC (Kilimanjaro)</td>
<td>1.12</td>
<td>0.50</td>
<td>.80</td>
</tr>
<tr>
<td>Meru DC (Arusha)</td>
<td>1.11</td>
<td>0.43</td>
<td>.79</td>
</tr>
<tr>
<td>Handeni DC (Tanga)</td>
<td>1.05</td>
<td>0.46</td>
<td>.92</td>
</tr>
<tr>
<td>Monduli DC (Arusha)</td>
<td>1.00</td>
<td>0.54</td>
<td>.99</td>
</tr>
<tr>
<td>Rombo DC (Kilimanjaro)</td>
<td>0.99</td>
<td>0.34</td>
<td>.97</td>
</tr>
<tr>
<td>Arusha DC (Arusha)</td>
<td>0.90</td>
<td>0.45</td>
<td>.84</td>
</tr>
<tr>
<td>Bumbuli DC (Tanga)</td>
<td>0.82</td>
<td>0.31</td>
<td>.61</td>
</tr>
<tr>
<td>Korogwe DC (Tanga)</td>
<td>0.70</td>
<td>0.23</td>
<td>.30</td>
</tr>
<tr>
<td>Lushoto DC (Tanga)</td>
<td>0.65</td>
<td>0.22</td>
<td>.20</td>
</tr>
<tr>
<td>Moshi MC (Kilimanjaro)</td>
<td>0.62</td>
<td>0.24</td>
<td>.22</td>
</tr>
<tr>
<td>Hai DC (Kilimanjaro)</td>
<td>0.58</td>
<td>0.20</td>
<td>.18</td>
</tr>
<tr>
<td>Mwanga DC (Kilimanjaro)</td>
<td>0.53</td>
<td>0.20</td>
<td>.10</td>
</tr>
<tr>
<td>Same DC (Kilimanjaro)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.51</td>
<td>0.17</td>
<td>.05</td>
</tr>
<tr>
<td>Muheza DC (Tanga)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.49</td>
<td>0.17</td>
<td>.04</td>
</tr>
<tr>
<td>Kilindi DC (Tanga)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.35</td>
<td>0.13</td>
<td>.005</td>
</tr>
<tr>
<td>Moshi DC (Kilimanjaro)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.29</td>
<td>0.09</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Abbreviations: CC, city council; DC, district council; EIR, electronic immunization registry; HCW, health care worker; MC, municipal council; TC, town council.

<sup>a</sup> Statistically significant at alpha=.05 level.

<sup>b</sup> Compared to Arusha city council, which was selected as it was the pilot implementation district and contains the capital and largest city in Arusha region.
tool, the odds of system use were much higher. In Tanzania, a national assessment in collaboration with the regional office in Tanga identified 33 facilities that transitioned to paperless reporting in March 2018; these facilities were selected to represent a mix of high and low performers with representation from all the district councils in Tanga. Since then (and not captured in our analysis), an additional 2 districts (60 facilities) in Tanga region transitioned to paperless reporting in September 2018. The Tanzania MOH and the President’s Office, Regional Administration and Local Government developed a checklist of criteria to inform the decision to transition facilities to paperless reporting. As more facilities migrate to paperless reporting, there is an opportunity to continue to test our hypothesis and confirm the results presented here that show paperless reporting increases system use.

In Zambia, those facilities that were farther from the DHO had significantly lower odds of using the system. Our initial hypothesis was that those farther facilities may have been less likely to have received supportive supervision or other district support due to their remote location. Distance from the DHO and other organizational covariates may have also captured dimensions of technical or behavioral facility characteristics. For instance, the distance from the DHO may capture farther facilities having limited infrastructure to support the technology, therefore lower likelihood of EIR use. (Indeed, distance to the DHO was collinear with urban/rural status in the Zambia model.) Similarly, in Tanzania, the increased likelihood of hospitals and health centers using the EIR may have also reflected that they were more likely to be connected to the electric grid.

### TABLE 5. Results of Regression Model Predicting EIR Use for Facilities in Zambia

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate (odds ratio)</th>
<th>Robust Standard Error</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Organizational</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 1 immunization day/week</td>
<td>0.82</td>
<td>0.31</td>
<td>.60</td>
</tr>
<tr>
<td>Distance from DHO, compared to 1st quartile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd quartile&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.46</td>
<td>0.15</td>
<td>.015</td>
</tr>
<tr>
<td>3rd quartile&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.41</td>
<td>0.14</td>
<td>.007</td>
</tr>
<tr>
<td>4th quartile&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.32</td>
<td>0.11</td>
<td>.001</td>
</tr>
<tr>
<td><strong>Behavioral</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weeks since EIR introduction&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.88</td>
<td>&lt;0.01</td>
<td>&lt;.001</td>
</tr>
<tr>
<td><strong>District&lt;sup&gt;b&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zimba</td>
<td>1.16</td>
<td>0.50</td>
<td>.74</td>
</tr>
<tr>
<td>Kazungula</td>
<td>0.97</td>
<td>0.48</td>
<td>.95</td>
</tr>
<tr>
<td>Mazabuka</td>
<td>0.91</td>
<td>0.40</td>
<td>.83</td>
</tr>
<tr>
<td>Livingstone</td>
<td>0.73</td>
<td>0.40</td>
<td>.57</td>
</tr>
<tr>
<td>Gwembe&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.33</td>
<td>0.17</td>
<td>.03</td>
</tr>
<tr>
<td>Kalomo&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.30</td>
<td>0.12</td>
<td>.003</td>
</tr>
<tr>
<td>Namwala&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.19</td>
<td>0.08</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Sinazongwe&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.14</td>
<td>0.08</td>
<td>.001</td>
</tr>
<tr>
<td>Pemba&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.10</td>
<td>0.04</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Monze&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.08</td>
<td>0.03</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Chikankata&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.05</td>
<td>0.03</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Siavonga&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.03</td>
<td>0.03</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Abbreviations: DHO, district health office; EIR, electronic immunization registry.
<sup>a</sup> Statistically significant at alpha=.05 level.
<sup>b</sup> Compared to Choma district, which was selected as it is the capital district for Southern Province and expected to be a high performer.
The technical covariates we tested were not statistically significantly associated with weekly EIR use in the final model for either country or were excluded due to missing data. Primary power source was significantly associated with weekly EIR use in Tanzania in bivariate analysis. However, facility type, public/private ownership, and the number of trained HCWs were associated with primary power missing data patterns, indicating the estimated association was biased. Primary power source was ultimately excluded from the final model due to missing data for Tanga region (Table 3). Although additional statistical methods would be needed to confirm the association between power supply and EIR use, the association observed in bivariate analysis was consistent with our hypothesis that facilities lacking a stable power source would be less likely to use the EIR. BID implementation staff in both countries had observed challenges with network availability, especially in rural facilities, or lack of data bundles that required some facilities to work in the EIR offline mode and later sync their data to the server.38

Finally, factors that were proxies for behavioral determinants—the number of HCWs initially trained and the number of weeks since EIR introduction—also emerged as important determinants. The number of weeks since EIR introduction was a strong predictor of decreased usage in both countries, with odds of system use declining by 1.9% each week in Tanzania and 12.4% in Zambia. This may have been due to waning HCW motivation or support to use the system over time. The more rapid decline in Zambia may have reflected the shorter window of ongoing support from the BID team post-EIR introduction compared to in Tanzania, the different intervention rollout strategies used,25 HCW perceptions of each country’s EIR system,12 or other factors. As mentioned, HCWs may have lost motivation to continue using the system if they perceived it as adding more work since they were still required to use the paper-based system for official reporting. Individual or facility recognition or incentives could have been powerful motivators to support behavior change but were applied inconsistently; failure to recognize consistent use combined with a lack of perceived data use may have contributed to waning motivation.

In Tanzania, the number of HCWs trained per facility during the EIR introduction period was significantly associated with increased odds of system use. This may suggest that facilities with more immunization HCWs had more bandwidth to use the system or perhaps that with more HCWs, they could have supported each other to continue system use. Across all facilities, additional support and accountability to encourage use of the EIR was needed to sustain use over time. In Zambia, the BID Initiative implementation phase ended shortly after EIR introduction in Southern Province was completed, so support significantly diminished. This may have, in part, explained the dramatic decline in system use over time. Demands on HCW time were numerous and without adequate incentives—whether financial or motivational—continuing a time-consuming practice may have been unreasonable to expect. Expectations of use from national or district leadership may have also played an important role in sustaining use.39 The MOHs in Tanzania and Zambia, in collaboration with the BID Initiative, continue to strengthen EIR usability. In Tanzania, there is ongoing work to identify additional facilities that are ready to transition to paperless reporting and to strengthen support for all facilities through a helpdesk line and district-level support. In Zambia, the MOH has focused on technical improvements to ZEIR, including fully integrating ZEIR with mVacc (an SMS-based platform to support community immunization awareness and access)40; transitioning data hosting to the MOH; refining district dashboards to allow for easy monitoring of facility data inputs; and reviewing potential syncing issues that may have impacted usage numbers reported in this study. In addition, an emphasis on engagement from local leaders at the subnational level has resulted in improved uptake of ZEIR among facilities. Like Tanzania, the Zambia MOH is assessing whether facilities can transition to paperless reporting to remove the burden of parallel systems. Finally, ZEIR is being used to monitor system use and data quality. The number of active users and the number of children present in the system each month are tracked as an indicator of use, and quality is assessed by comparing EIR records against those in paper registries.

This article presents one way of measuring use—based on the dates of data entered into the EIR—but there are other ways that system use could be measured. We chose this measure because it could be consistently measured across the available data from TImR and ZEIR. Using the dates attached to when services were delivered also allowed us to compare system use across facilities working in the online and offline modes; for example, if a facility entered a vaccine
Facility-Level Use of Electronic Immunization Registries in Tanzania and Zambia

As more countries move to introduce EIRs or other digital interventions, we recommend that they measure and monitor use of the system(s) among the intended users. New systems that aim to improve data timeliness, availability, or completeness will only be able to do so if they are used consistently as intended. Indicators for monitoring system use through metadata have been published by the World Health Organization and can be integrated into program plans. Measuring system use is important to: (1) inform interpretation of the data, since traditional reporting measures like coverage and dropout rates may be skewed depending on the completeness of data entry; and (2) inform programmatic decisions, such as targeting support to facilities with suboptimal use or determining when facilities are ready to transition away from traditional paper-based tools to using the EIR as their primary data collection and reporting tool.

**Limitations**

A key limitation of the analysis was data unavailability. Most available secondary data captured organizational and technical factors. We were limited in our ability to measure the impact of behavioral factors, including individual attitudes, skills, and motivation, which may be important in explaining differences in use across facilities. Also, some covariates had too much missing data to include in our models, such as primary power source in Tanzania. As a result, determinants included in the models may have been impacted by unmeasured confounding. For example, greater odds of EIR use at hospitals and health centers compared to dispensaries may have, in part, reflected greater access to electricity and internet. For the covariates we did include, most did not have information available on how they have changed over time, which may limit their ability to explain changes in use over time. For example, we captured the number of HCWs trained at the point of EIR introduction, but we do not have information on staff turnover. Finally, our analysis used all available EIR data for each region and given different EIR introduction dates, this resulted in a different number of data points per district. In Tanzania, trends may be driven by more data points from Arusha region compared to more limited follow-up time in regions that introduced the EIR later.

**CONCLUSION**

The results from this analysis add to our understanding of the organizational and behavioral factors associated with facility EIR use in Tanzania and Zambia. This analysis demonstrated greater EIR utilization among facilities at which HCWs reported into electronic systems only compared to parallel paper and electronic systems. In addition, it highlighted the importance of ongoing support for new digital interventions during and beyond the initial rollout, as demonstrated by greater EIR utilization at facilities with a greater number of HCWs trained in the intervention and at facilities closer to the DHO, and by decreased utilization over time. Strong district leadership and the mentorship and close supervision of HCWs have been considered essential to successful uptake and ongoing use of these systems.

The MOH and BID Initiative teams in each country should continue to collaboratively identify ways to improve EIR system use, such as continuing to transition facilities to paperless reporting, promoting the benefits of system use to HCWs, and making system use metrics available (e.g., through automated reports or dashboards) to empower stakeholders at all levels to monitor and support consistent use. System use metrics may need to be triangulated with other data sources (e.g., HCW surveys) or evaluation approaches to explain the observed trends in use.

Stakeholders introducing EIRs, or other digital health interventions, should consider providing additional support to more remote, lower-volume facilities and should develop plans from the start for when and how to transition to paperless reporting. However, these factors associated with system use may vary in different contexts. As EIRs are introduced in new contexts, we recommend building...
these types of analyses directly into the system. Moreover, if EIRs are designed to capture more granular information about facility characteristics and/or are linked to other routine health information systems, then additional data can be available to understand the different factors influencing system use.

Acknowledgments: The authors would like to thank the Government of Tanzania and the Government of Zambia, and specifically the immunization programs, for their deep thought partnership in developing these interventions and evaluating their efficacy. This includes Dr. Dalrossa Lyimo, Dr. Mpkapi Ullusubia, Hermes Satter, and Dr. Ngwegwe Bulula in Tanzania, and Dr. Francis Mwansa and Brivine Sikapande in Zambia. In addition, the authors would like to acknowledge the contributions of additional members of the analytics team, including Samantha B. Dolan, Jonathan Drummey, Sarah Skye Gilbert, Emma Korpi, Hil Lyans, Hassan Mttenga, Fred Nzabu, Chilungu Pula, John Richard, and Jenny Thompson. Finally, the authors would like to thank the Bill & Melinda Gates Foundation for supporting this research, and particularly thank Tove Ryman for her insights and guidance.

Funding: The authors acknowledge funding from the Bill & Melinda Gates Foundation to support this work.

Conflicting interests: None declared.

REFERENCES

26. Aqil A, Lippeveld T, Haozumi D. PRISM framework: a paradigm shift for designing, strengthening and evaluating routine health...
CrossRef. Medline


Peer Reviewed

Received: March 30, 2020; Accepted: July 7, 2020; First published online: August 19, 2020


© Carnahan et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are properly cited. To view a copy of the license, visit http://creativecommons.org/licenses/by/4.0/. When linking to this article, please use the following permanent link: https://doi.org/10.9745/GHSP-D-20-00134

Facility-Level Use of Electronic Immunization Registries in Tanzania and Zambia www.ghspjournal.org

Global Health: Science and Practice 2020 | Volume 8 | Number 3

504
The All Babies Count Initiative: Impact of a Health System Improvement Approach on Neonatal Care and Outcomes in Rwanda

Hema Magge, a,b,c Evrard Nahimana, c Jean Claude Mugungu, d Fulgence Nkikabahizi, e Elisabeth Tadiri, f Felix Sayinzoga, e Anatole Manzi, d Merab Nyishime, c Francois Biziyaremye, c Hari Iyer, g Bethany Hedt-Gauthier, h Lisa R. Hirschhorn i

Key Findings

- A district-wide health system improvement program combining facility readiness support, clinical training/mentoring, and district-wide improvement collaboratives increased quality improvement capacity, improved maternal and newborn quality of care, and reduced neonatal mortality by approximately 35% overall and 49% among high-risk preterm/low birth weight infants.
- This improvement in mortality was not seen during the same time period in the rest of rural Rwanda.

Key Implications

- Policy makers should consider adopting components of the All Babies Count program into the design of system improvement approaches to transform quality of care and outcomes for newborns.
- Embedding the design into existing health system structures could help structure improvement in other clinical domains.

ABSTRACT

Introduction: Poor-quality care contributes to a significant portion of neonatal deaths globally. The All Babies Count (ABC) initiative was an 18-month district-wide approach designed to improve clinical and system performance across 2 rural Rwandan districts.

Methods: This pre-post intervention study measured change in maternal and newborn health (MNH) quality of care and neonatal mortality. Data from the facility and community health management information system and newly introduced indicators were extracted from facility registers. Medians and interquartile ranges were calculated for the health facility to assess changes over time, and a mixed-effects logistic regression model was created for neonatal mortality. A difference-in-differences analysis was conducted to compare the change in district neonatal mortality with the rest of rural Rwanda.

Results: Improvements were seen in multiple measures of facility readiness and MNH quality of care, including antenatal care coverage, preterm labor management, and postnatal care quality. District hospital case fatality decreased, with a statistically significant reduction in district neonatal mortality (odds ratio [OR]=0.54; 95% confidence interval [CI]=0.36, 0.83) and among preterm/low birth weight neonates (OR=0.47; 95% CI=0.25, 0.90). Neonatal mortality was reduced from 30.1 to 19.6 deaths/1,000 live births in the intervention districts and remained relatively stable in the rest of rural Rwanda (difference in differences =12.9).

Conclusion: The ABC initiative contributed to improved MNH quality of care and outcomes in rural Rwanda. A combined clinical and health system improvement approach could be an effective strategy to improve quality and reduce neonatal mortality.

INTRODUCTION

Significant progress has been made worldwide in under-5 mortality, with more than 50% reduction from 1990 to 2015. However, the rate of progress in the neonatal period has been slower, with death in the first month of life accounting for 46% of under-5 deaths globally. In response, a global movement has grown to accelerate progress, and efforts have prioritized improving access along antenatal, delivery, and postnatal care. However, interventions targeting access without addressing quality have not resulted in reduction of...
neonatal mortality; improving care quality could prevent up to an estimate 71% of newborn deaths.

After the health system devastation from the 1994 genocide, Rwanda rebuilt the health sector, experienced a 63% reduction in premature mortality between 2000 and 2011, and achieved Millennium Development Goals 4 and 5. However, similar to the global experience, neonatal mortality reduction lagged behind the success in reducing deaths after the first month of life. In response, Rwanda prioritized addressing neonatal mortality. By 2010, Rwanda had tremendous increases in facility-based maternal and newborn health (MNH) care including skilled delivery; therefore, improving facility newborn care quality became a key priority.

Since 2005, Partners In Health (PIH), a global nongovernmental organization with a mission to provide high-quality health services in disadvantaged communities, has supported the Rwanda Ministry of Health (MOH) in district health system strengthening. In response to the prioritization of reducing neonatal mortality nationally, PIH partnered with MOH to design the All Babies Count (ABC) initiative to improve district-wide quality of newborn care. The objective of this study was to evaluate the impact of ABC on quality improvement (QI) activities, neonatal quality of care, and neonatal mortality across the prenatal, perinatal, and postnatal risk periods in 2 rural Rwandan districts.

**METHODS**

**Setting**

This study was conducted in southern Kayonza (SK) and Kirhe districts in eastern Rwanda, where PIH began support in 2005. The MOH chose these districts in part due to their poor health status post-genocide. The intensity of PIH support was gradually reduced as the 2 districts caught up with the rest of the country in infrastructure, economy, and health status. By 2012, PIH’s partnership had moved from direct care provision and management support toward targeted financial and technical assistance to support clinical innovation areas, including newborn health.

ABC was introduced in 2013 in SK and Kirhe districts, with a population of approximately 500,000 people. At initiation, each district had 1 hospital, with 8 health centers providing maternity services in SK and 13 in Kirhe. These government facilities were administered and funded publicly and received additional technical/financial support for general operations and specialized clinical services through their partnership with PIH. Three additional health centers initiated maternity services in Kirhe during ABC implementation and were integrated into the program. District hospitals were staffed per national standards by general practitioners, nurses, and midwives, and they had the capacity to perform cesarean deliveries. Each district hospital had a neonatal unit as per the national standards to provide care for ill and preterm infants using the national neonatal care protocol including continuous positive airway pressure management. A US-trained pediatrician supported by PIH was present 6 months per year as part of general hospital support across the 2 districts. Health centers were staffed by nurses and midwives performing routine MNH and under-5 care.

**Program Design**

Building upon a mentoring-based QI program that was successful in addressing quality gaps in pediatric care across the 2 districts, ABC was designed to reduce neonatal mortality by improving the quality of antenatal, delivery, and postnatal care through combining facility-focused clinical and QI mentorship with district-wide QI collaboratives adapted from the Breakthrough Series model (Figure). Our aim was to ensure facility readiness for quality newborn care at the start, followed by the introduction of a district-wide QI approach to address key mortality drivers along the continuum of antenatal, delivery management, and postnatal care and at all levels of the system. The approach was designed in collaboration with the MOH Maternal and Child Health Department, with feedback from key national and district stakeholders.

A baseline quality assessment was completed by the PIH monitoring and evaluation team as part of a routine quarterly health facility survey. By design, the results were used by the ABC team to identify and close gaps in relevant neonatal training and commodities before the first learning session (LS). Gaps were addressed through baseline provision of essential equipment to meet facility national standards and annual relevant clinical trainings to address staff turnover. The QI collaboratives aimed to strengthen health care worker QI and clinical capacity through mentorship; efficiently test a large number of locally designed system interventions, called “change ideas,” to create relevant neonatal change packages; and spread successful changes through LSs. Each collaborative was based on district leadership and interdisciplinary QI teams from all district facilities (hospital.
and health centers), including facility directors, midwives, antenatal/postnatal care providers, data officers, and community health supervisors. Hospital QI teams included the clinical director and/or head of maternity, neonatal, and operating theaters. On average, 3–5 team members participated in each LS (total 30–50 participants).

PIH-employed ABC nurse mentors were neonatal clinical experts trained in QI, data quality, and mentorship. They provided the integrated clinical and QI mentorship and supported QI projects to all district facilities (hospital and health centers). They worked alongside district hospital supervisors—whose existing responsibilities included supervising clinical care at health centers—to build their technical and mentorship skills to transition program ownership by the end of the collaborative. ABC mentors and district hospital supervisors worked together on core program implementation elements including data collection, facility mentorship, and LS organization and facilitation.

Study Design
The program evaluation was a pre-post intervention design measuring change in MNH quality of care and neonatal mortality. Embedded into the study were measurements of program implementation, QI activities, facility readiness, and patient satisfaction. Stillbirths were tracked as a secondary outcome. Baseline was July–September 2013 with the first LSs in October 2013. Because neonatal mortality had seasonal variability, the endpoint quarter was chosen 3 months after the last LS to capture the same season (July–September 2015). Indicators that were newly introduced had a baseline of October–December 2013 (Table 1).

Measures
Program Implementation
We measured the implementation process and outcomes including domains of feasibility and fidelity based on the framework published by Proctor et al. The data collected included trainings and equipment provided, mentoring frequency and content, LS attendance, and change ideas tested by facilities. These data were used to strengthen program delivery and allow for adaptations to maximize impact.

QI Activities and Facility Readiness
The measurement of facility QI activities included the number of projects executed and change ideas tested. Facility readiness was based on national facility neonatal standards including minimum staffing, equipment, and medications (Supplement 1).

MNH Quality of Care and Neonatal Mortality
MNH quality indicators were selected based on literature review, MOH priorities, and existing health management information system (HMIS) data to represent the evidence-based care pathways: antenatal, delivery management, and postnatal care (Table 1). Measurement of asphyxia and measurement of provision of antibiotics for
**TABLE 1. Selection of Maternal and Newborn Health Core Improvement Collaborative Indicators and Outcome Measures Used in Evaluating an All Babies Count Initiative Implemented in Rwanda**

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Numerator</th>
<th>Denominator</th>
<th>Data Source</th>
<th>Facility</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antenatal care</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1. ANC standard 4 visit coverage</td>
<td>Number of women who delivered in a HF who had 4 ANC visits</td>
<td>Total number of health facility deliveries PLUS total number of referrals of laboring women from HC to hospital PLUS home deliveries</td>
<td>HMIS</td>
</tr>
<tr>
<td><strong>Delivery services</strong></td>
<td>2. Percentage of pregnant women with facility delivery</td>
<td>Total number of health facility deliveries</td>
<td>Total number of home deliveries PLUS total number of health facility deliveries</td>
<td>Community HMIS</td>
</tr>
<tr>
<td></td>
<td>3. Time to cesarean delivery for emergency (from determination of need at hospital to time of delivery)</td>
<td>Average time to cesarean delivery for emergency (from determination of need by doctor at hospital to time of cesarean delivery incision)</td>
<td>Chart review&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Hospital</td>
</tr>
<tr>
<td></td>
<td>4. Antenatal steroids for preterm labor</td>
<td>Number of women with preterm labor (&lt;34 weeks treated with dexamethasone)</td>
<td>Number of women with preterm labor (&lt;34 weeks)</td>
<td>Chart review</td>
</tr>
<tr>
<td><strong>Postnatal care</strong></td>
<td>5. Percentage of babies with immediate skin-to-skin after delivery</td>
<td>Number of babies placed immediately skin to skin</td>
<td>Number of babies born vaginally</td>
<td>Chart review</td>
</tr>
<tr>
<td></td>
<td>6. Danger sign assessment within 24 hours</td>
<td>Newborns checked for danger signs in postpartum ward within 24 hours</td>
<td>Total deliveries MINUS stillbirths macerated MINUS stillbirths fresh MINUS death at birth of live born babies</td>
<td>HMIS</td>
</tr>
<tr>
<td><strong>Outcome</strong></td>
<td>7. Neonatal unit case fatality</td>
<td>Number of deaths in neonatal unit (&lt;28 days)</td>
<td>Number of admissions to neonatal unit (&lt;28 days)</td>
<td>Chart review</td>
</tr>
<tr>
<td></td>
<td>8. District-wide neonatal mortality</td>
<td>HMIS neonatal deaths PLUS death at birth of live born babies PLUS number of neonatal deaths at community</td>
<td>Live births PLUS number of home deliveries</td>
<td>HMIS</td>
</tr>
<tr>
<td></td>
<td>9. Facility neonatal mortality among preterm/LBW&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Hospital deaths due to prematurity PLUS HC deaths due to prematurity</td>
<td>Hospital LBW/non-preterm PLUS hospital preterm PLUS HC LBW/non-preterm PLUS HC preterm</td>
<td>HMIS</td>
</tr>
<tr>
<td></td>
<td>10. Facility neonatal mortality among non-preterm/LBW&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Hospital deaths of all causes except prematurity PLUS HC deaths of all causes except prematurity</td>
<td>Hospital and HC live births MINUS (hospital LBW/non-preterm PLUS hospital preterm PLUS HC LBW/non-preterm PLUS HC preterm)</td>
<td>HMIS</td>
</tr>
<tr>
<td></td>
<td>11. Facility stillbirths (macerated and fresh)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Stillbirths macerated PLUS stillbirths fresh</td>
<td>Total deliveries</td>
<td>HMIS</td>
</tr>
</tbody>
</table>

Abbreviations: ANC, antenatal care; HC, health center; HF, health facility; HMIS, health management information system; LBW, low birth weight.

<sup>a</sup>Rwinkwavu health center did not have delivery services, so the denominator was changed to number of community HMIS health facility deliveries at district hospital plus number of community HMIS home deliveries.

<sup>b</sup>Chart review indicators were newly introduced at the start of the intervention.

<sup>c</sup>Prematurity defined as gestational age ≤37 weeks, and low birth weight defined as birth weight < 2500 g per World Health Organization standard definition.

<sup>d</sup>Stillbirths defined in national HMIS data dictionary as a baby born with no signs of life at or after 22 weeks gestation and with birth weight greater than or equal to 500 g. Stillbirth analysis restricted to facility level given the absence of community-based recording of stillbirths.
premature preterm rupture of membranes were introduced, but reliable capture was found to be a challenge for providers due to difficulty with systematic documentation. Despite being unable to reliably track the number of newborns with asphyxia, asphyxia prevention and management was a core focus of clinical mentorship and addressed through many QI change ideas targeting an aspect of delivery management. Hospital neonatal unit and district-wide neonatal mortality was measured using all available data (including births and deaths in the community). Mortality among preterm/low birth weight (LBW) infants was estimated using available data (facility only) as a particularly high-risk subpopulation and a leading cause of neonatal mortality.

Data Collection
Implementation data and QI activities were extracted from routine program tools. Measurement of essential equipment, medications, and training standards were assessed quarterly using standardized service readiness surveys.

MNH process and mortality data from facility HMIS were extracted and compared with paper registers as part of QI coaching. Data for new indicators were collected directly from additional facility registers and through weekly random sampling of charts for cesarean delivery time. Births and deaths in the community were collected from community HMIS and compared with community health worker supervision records from the corresponding facilities. District neonatal mortality for the rest of rural Rwanda was constructed from the HMIS as register review was not feasible.

Patient satisfaction data were collected using surveys measuring experience of care and satisfaction on a Likert scale from 1 (poor) to 5 (excellent) of a sample of women attending antenatal care (ANC) and delivery services in all intervention facilities. This study included 278 women from Kirehe (ANC 204, maternity 74) and 198 from SK (ANC 166, maternity 32) with baseline collection from November to December 2013, and endpoint collection from July to September 2015. Detailed methods are described elsewhere.19

Analysis
Program Implementation, MNH Quality of Care, and Neonatal Unit Case Fatality
We constructed the numerator, and where appropriate the denominator, by calculating the mean monthly value for the 3-month period corresponding to the baseline or endpoint quarter. For each time period, we report the median of the numerator and the median of the denominator. For indicators 1–4, 6, 8, and 9 in Table 2, we report the median and interquartile range for the health facility and assessed changes using a Wilcoxon signed rank test at the α=.05 significance level. Analyses were conducted in Stata SE v14 (College Station, TX).

Patient Satisfaction
We assessed the difference in the proportion reporting high (Likert score=4 or 5) patient satisfaction between baseline and endpoint using a chi-squared test at the α=.05 significance level.

Neonatal Mortality and Stillbirths
To estimate neonatal mortality at the level of the individual birth or death rather than as a population estimate aggregated to facility level, we expanded the dataset such that each row corresponded to either a neonatal death or live birth. To offset the power increase obtained by artificially increasing the number of observations, we included a categorical variable for health facility to use as a random intercept and account for clustering. This pseudo-dataset contained the following variables: indicator for death (1=death/0=live birth), time period (1=baseline, 0=post), health facility (categorical for each health facility), and district (1=SK, 0=Kirehe). Based on our hypothesis that mortality could differ based on gestational age/birth weight, mortality estimates were calculated for each subgroup. For each mortality measure, we used mixed-effects logistic regression models to assess changes in mortality with a random effect for health facility to account for clustering (Supplement 2). We tested for interactions between intervention and district with an intervention-district interaction term included in our regression model. If the interaction term was statistically significant, we reported the intervention effect stratified by district. When the interaction term was not significant, we reported a collapsed effect (Supplement 3). The models were fit using SAS v 9.4 and did not control for other covariates because they were not available on the individual level.

Neonatal Mortality Difference in Differences
Change in mortality was defined as the difference of population deaths per 1,000 live births between baseline and endpoint. The difference in differences was examined for the intervention districts against HMIS-reported data from the rest of Rwanda (24 districts with 8.4 million people).
including urban Kigali and Burera District). Burera District was excluded because it received PIH neonatal clinical and infrastructure support, but not the complete ABC program. Population district mortality for the national comparison districts included all HMIS mortality data available from the relevant quarters.

**Ethical Considerations**

Informed consent was obtained from women surveyed for patient satisfaction.19 This study was approved by the Institutional Review Board of Brigham and Women’s Hospital (2009-P-001941/11; BWH) and the Rwanda National Ethics Committee (RNEC 032/RNEC/2012).

**Patient and Public Involvement**

The study was supported by a community advisory group composed of separate focus groups with women from the intervention area, community health workers, traditional healers, and facility nurses and doctors. The advisory group provided input for the initial program conceptual framework and intervention design. They were not

### TABLE 2. Change in Maternal and Newborn Health Quality of Care Indicators in All Babies Count Initiative Implemented in 2 Districts in Rwanda

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Aggregate</th>
<th>South Kayonza</th>
<th>Kirehe</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>Endpoint</td>
<td>Median (IQR)</td>
</tr>
<tr>
<td>1. Percentage of deliveries where mothers had 4 standard ANC visits</td>
<td>13.7</td>
<td>30.4</td>
<td>(6.7, 44.2)</td>
</tr>
<tr>
<td>2. Percentage of pregnant women delivering in facilities</td>
<td>89.6</td>
<td>92.6</td>
<td>(86.3, 94.9)</td>
</tr>
<tr>
<td>3. Percentage of babies who are provided immediate skin-to-skin after birth</td>
<td>53.6</td>
<td>97.4</td>
<td>(0, 80.9)</td>
</tr>
<tr>
<td>4. Percentage of newborns checked for danger signs within 24 hours of birth</td>
<td>46.6</td>
<td>98.7</td>
<td>(31.1, 96.7)</td>
</tr>
<tr>
<td>5. Average hospital time to emergency cesarean delivery (minutes)</td>
<td>167</td>
<td>50</td>
<td>–</td>
</tr>
<tr>
<td>6. Percentage of women with preterm labor who are treated with antenatal steroids</td>
<td>0</td>
<td>41.7</td>
<td>(0, 100)</td>
</tr>
<tr>
<td>7. Percentage of facilities with at least 2 MNH clinically trained staff</td>
<td>100</td>
<td>100</td>
<td>–</td>
</tr>
<tr>
<td>8. Percent availability of essential medications for MNH care</td>
<td>61.2</td>
<td>81.8</td>
<td>(45.0, 77.8)</td>
</tr>
<tr>
<td>9. Percent availability of functioning equipment essential for MNH care</td>
<td>55.6</td>
<td>86.6</td>
<td>(48.2, 61.1)</td>
</tr>
<tr>
<td>10. Patient satisfaction*: average satisfaction with ANC</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>11. Patient satisfaction: average satisfaction with maternity care</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Abbreviations: ANC, antenatal care; IQR, interquartile range; MNH, maternal and newborn health; SD, standard deviation.

*Patient satisfaction scores on a Likert scale: 1=excellent; 2=very good; 3=good; 4=fair; 5=poor.
specifically involved in the evaluation methods design because the evaluation was predominantly conducted using routine programmatic data. Findings from the study have been disseminated with key stakeholders within the health system and communities affected by the work.

Role of the Funding Source
The funder was not involved in study design, execution, or preparation of this manuscript.

RESULTS
Program Implementation and QI Activities
The ABC initiative was implemented in 2 districts and reached all facilities, with full participation by district leadership and the facility QI teams. There was no significant change in clinical staffing of health centers (baseline 9, endpoint 10; \( P = .09 \)). Fidelity include mentoring coverage. Facilities received an average of 0.68 visits/month, and a change package of 46 successful change ideas was developed by the endpoint to facilitate spread.16 ABC adaptation informed by quarterly implementation data was undertaken to increase acceptability and adoption. For example, initially LSs were planned quarterly to complete the collaborative within 1 year. However, during LS2, it was found that more time was needed for data quality stabilization, introduction of data sources for new indicators (or HMIS indicators without clear register sources), QI project implementation, and data monitoring and coaching activities. In response, the collaborative duration was adapted to 18 months to extend action periods to accommodate having adequate time for change idea testing. Adoption of ABC components by facilities was seen, including QI (with a median of 2 QI projects running at endpoint, a total of 118 change ideas were tested across the 2 collaboratives), and complete facility attendance at all LSs in both districts.

Facility Readiness, MNH Quality of Care, and Mortality
Table 2 shows changes in quality indicators that were the locus of QI activities. Facility readiness improved significantly, including availability of essential medications (median difference=15.1%; interquartile range [IQR]=4%, 29.8%; \( P < .001 \)) and equipment (median difference=31.2%; IQR=−19.5%, 37.8%; \( P < .001 \)). As planned, clinical trainings were conducted before the baseline quality of care data collection; all facilities had at least 2 MNH-trained staff, Rwinkwavu district hospital had 8 neonatal intensive care unit-trained staff at baseline and endpoint, and Kirehe increased from 7 to 14. The specific individuals may have changed. Significant improvement in ANC coverage (median difference=13.6%; IQR=−3.3%, 26.4%; \( P = .04 \)), provision of immediate skin-to-skin (median difference=43.6%; IQR=17.7%, 95.7%; \( P < .001 \)), and danger signs assessment (median difference=47.7%; IQR=−1.4%, 67.1%; \( P < .001 \)), was seen across the intervention districts. Results in these measures improved across both districts with the exceptions of ANC coverage and immediate skin-to-skin, which had higher baselines in SK. Facility delivery rate had baseline values approaching 90% in both districts, and rates were sustained (median difference=7.8%; IQR=−2.8%, 5.1%; \( P = .25 \)). District-level improvement was seen in complications management: steroid administration for preterm labor increased from 0 to 41.7% (median difference=0%; IQR=0%, 100%; \( P = .32 \) across both) and time from cesarean delivery decision to incision decreased in SK from 167 to 50 minutes and in Kirehe from 82 to 61 minutes). District hospital neonatal unit case fatality decreased from 28.2% to 12.2% in SK and from 23.4% to 10.1% in Kirehe.

Patient-reported satisfaction with care for both ANC and maternity services had minimal change. For ANC, no significant change was seen in SK (from 2.8 to 3.1; \( P = .11 \)) and minimal improvement was seen in Kirehe (from 3.3 to 3.6; \( P = .01 \)). No change was seen in either district for satisfaction with maternity care.

Table 3 presents results from the mixed-effects logistic regression models estimating change in neonatal mortality associated with ABC implementation. District neonatal mortality significantly decreased overall from 30.1 to 19.6 deaths/1,000 live births (adjusted odds ratio [aOR]=0.54; 95% confidence interval [CI]=0.36, 0.83). Among preterm/LBW neonates, mortality decreased from 198.8 to 100.6 deaths/1,000 preterm/LBW live births (aOR=0.47; 95% CI=0.25, 0.90). Mortality among non-preterm infants had a nonsignificant decrease from 10.4 to 7.5 deaths/1,000 non-preterm/LBW live births (aOR=0.60; 95% CI=0.36, 1.02). Stillbirths were the only outcome for which district was found to be a significant effect in the model and are reported by district. A nonsignificant decrease was found in Kirehe (OR=0.90; 95% CI=0.61, 1.32), and there was a significant increase in SK (OR=1.71; 95% CI=1.06, 2.75).

District neonatal morality significantly decreased overall and district hospital neonatal unit fatality decreased.

Significant improvement in ANC coverage, provision of immediate skin-to-skin contact, and danger signs assessment was seen.
Neonatal mortality decreased by 45% in the intervention districts, while remaining relatively stable in the rest of rural Rwanda (Table 4). The difference in differences analysis found a notable difference in the change in the intervention area compared with national secular trends (−13.0).

Given the lack of register-reported data in the comparison districts, a sensitivity analysis was conducted comparing the HMIS-reported data from the intervention districts, and similarly found a notable difference in differences of −9.2.

### DISCUSSION

**Implementation Successes**

We found that the successful implementation of a multilevel intervention combining facility readiness, clinical mentoring, and district-wide improvement collaboratives increased QI capacity, improved quality of care, and was temporally associated with reduced neonatal mortality overall and among preterm/LBW infants—a high-risk subpopulation—in the intervention districts. Maternal and neonatal quality of care in low- and middle-income countries has gained attention as increases in service coverage have not been met with anticipated mortality reduction. Attention is now drawn to health system improvement as a strategy to improve outcomes.

We approached neonatal mortality reduction as a district-wide endeavor—a factor that we believe facilitated impact. We aimed to align facility readiness and care provision with national standards and strengthen quality across all facilities acting as a network of care to serve a catchment population. We leveraged the standard improvement collaborative approach—which typically focuses on facilities through a defined learning network—to create a district-wide learning system to accelerate improvement.

Although improvement collaboratives have gained popularity based on success in high-income countries, our approach—combining multiple levels—provided a comprehensive solution. This strategy could be replicated in similar settings to achieve significant reductions in neonatal mortality.

### TABLE 3. Change in Neonatal Mortality Across 2 Districts in Rwanda Where All Babies Count Initiative Was Implemented

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Baseline</th>
<th>Endpoint</th>
<th>aOR (95% CI)a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neonatal mortality (deaths/1,000 live births)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggregate</td>
<td>30.1 (103/3,426)</td>
<td>19.6 (57/2,902)</td>
<td>0.54 (0.36, 0.83)</td>
</tr>
<tr>
<td>Southern Kayonza</td>
<td>35.4 (55/1,553)</td>
<td>18.5 (24/1,295)</td>
<td></td>
</tr>
<tr>
<td>Kirehe</td>
<td>25.6 (48/1,873)</td>
<td>20.5 (33/1,607)</td>
<td></td>
</tr>
<tr>
<td>Facility neonatal deaths in preterm infants,1,000 preterm and LBW live births</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggregate</td>
<td>198.8 (32/161)</td>
<td>100.6 (18/179)</td>
<td>0.47 (0.25, 0.90)</td>
</tr>
<tr>
<td>Southern Kayonza</td>
<td>290.3 (18/62)</td>
<td>134.3 (9/67)</td>
<td></td>
</tr>
<tr>
<td>Kirehe</td>
<td>141.4 (14/99)</td>
<td>80.4 (9/112)</td>
<td></td>
</tr>
<tr>
<td>Facility neonatal deaths in non-preterm babies,1,000 non-preterm and LBW live births in district</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggregate</td>
<td>10.4 (36/3,446)</td>
<td>7.5 (24/3,181)</td>
<td>0.60 (0.36, 1.02)</td>
</tr>
<tr>
<td>Southern Kayonza</td>
<td>7.9 (11/1,387)</td>
<td>5.9 (7/1,177)</td>
<td></td>
</tr>
<tr>
<td>Kirehe</td>
<td>12.1 (25/2,059)</td>
<td>8.5 (17/2,004)</td>
<td></td>
</tr>
<tr>
<td>Facility stillborn rate (macerated and fresh)/total per 1,000 births</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggregate</td>
<td>23.4 (84/3,590)</td>
<td>28.8 (99/3,436)</td>
<td></td>
</tr>
<tr>
<td>Southern Kayonza</td>
<td>20.7 (29/1,398)</td>
<td>34.5 (44/1,274)</td>
<td>1.71 (1.06, 2.75)</td>
</tr>
<tr>
<td>Kirehe</td>
<td>25.1 (55/2,192)</td>
<td>25.4 (55/2,162)</td>
<td>0.90 (0.61, 1.32)</td>
</tr>
</tbody>
</table>

Abbreviations: aOR, adjusted odds ratio; CI, confidence interval; LBW, low birth weight.

aOdds estimates from models stratified by district only provided if test for interaction between district and time in regression was statistically significant at 0.05 level.
settings, evidence of impact in low- and middle-income countries has been mixed. However, to our knowledge, evidence has not been published regarding collaboratives in low- and middle-income countries adapted to facilitate a district-wide health system approach. We used LSs to convene key stakeholders, from community representatives to district leaders, to facilitate problem solving across traditional hierarchies. MNH quality gaps that require multilevel solutions, such as infrastructure, referral systems, and data quality, were tackled in improvement projects. The ABC initiative, as with all QI efforts, required ownership by leaders to achieve impact. We achieved this through intensive and proactive stakeholder engagement. It included partnering with national and district leadership to determine what they believed would be necessary to achieve high-quality MNH care, ensure district hospital capacity to manage neonatal complications, and create functioning district-wide referral systems. Evidence of leadership engagement was seen through the prioritization of addressing newborn health in a coordinated manner. Both districts’ leaders included neonatal quality of care and mortality into their public performance contracts (imihigo) to which they were formally accountable to the government, and they led efforts to improve the culture of data reporting and use. We found the inclusion of MNH goals in the district leadership imihigo to be a way to help engage leaders throughout the duration of program implementation and to jointly solve challenges encountered. We saw reduced discrepancies in the HMIS of neonatal deaths over time, indicating some degree of success of these efforts.

Within the collaboratives, we saw broad QI adoption reflected in activated facility teams with QI capability and a high degree of QI activities across facilities. QI is often found to be focused on microlevel facility improvements, resource intensive, and of questionable impact; however, we found this district-wide system approach feasible and effective.

### Implementation Challenges

Furthermore, the QI work endured despite challenges. Illustrative examples include staff turnover, a data culture of “blame,” and changing population needs. First, staff at health centers moved frequently—as is commonly faced in low- and middle-income countries. Although overall staff number remained stable, changes in individuals required an adaptive process to bring new staff in QI teams up to date. We found the improvement collaborative design, when integrated into routine structures, provided a scaffolding for quickly bringing new staff into facility and district

| TABLE 4. Change in Neonatal Mortality in All Babies Count Intervention Area in 2 Districts Compared With the Rest of Rural Rwanda |
|---------------------------------|---------------------------------|---------------------------------|
|                                 | Pre-intervention (Deaths/1,000 Live Births) | Post-intervention (Deaths/1,000 Live Births) | Per 1,000 Change |
| (1) Southern Kayonza/Kirehe (ABC) | 30.1 (103/3,426) | 19.6 (57/2,902) | −10.4 |
| Southern Kayonza                | 35.4 (55/1,553)  | 18.5 (24/1,295) |   |
| Kirehe                          | 25.6 (48/1,873)  | 20.5 (33/1,607) |   |
| (2) Southern Kayonza/Kirehe (HMIS) | 22.3 (87/3,896) | 15.6 (59/3,785) | −6.7 |
| Southern Kayonza                | 28.7              | 25.3              |   |
| Kirehe                          | 18.3              | 10.7              |   |
| (3) Rural Rwanda comparison districts⁹ | 13.4 (834/62,382) | 15.9 (958/60,225) | 2.5 |

Abbreviations: ABC, All Babies Count; HMIS, health management information system.

⁹ Bugesera, Gakenke, Gatsibo, Gicumbi, Gisagara, Huye, Kamonyi, Karongi, Muhanga, Musanze, Ngoma, Ngoroero, Nyabihu, Nyagatare, Nyamagabe, Nyamasheke, Nyanza, Nyaruguru, Rubavu, Ruhango, Rulindo, Rusizi, Rutsiro, Rwamagana Districts HMIS reported data.
QI efforts and enabling them to be clinically mentored and learn QI methods.

Another challenge encountered early in the program related to leadership support at health centers. Leaders were engaged in stakeholder meetings; however, once QI teams began working to address system problems, some participants reported resistance from their supervisors attributed to lack of familiarity with QI methods. As in many settings, creating a culture of data use requires shifting from norms of blaming individuals, to norms of understanding and improving systems. Therefore, mentors and district supervisors trained district and facility leaders in QI methods and included these leaders in LSs to build their QI capability and data fluency.

Contextual changes in the population and health system also posed a challenge. In Kirehe, given the large population size of some facility catchment areas, 3 new health centers were built during the initiative and had to be incorporated into coaching and LSs. Additionally, the district received a large influx of returning Rwandans from Tanzania in 2013. The improvement collaborative design allowed for natural integration of new facilities into LSs, coaching visits, and QI initiatives, as well as customized support to respond to these changes.

More generally, the multilevel design facilitated system gap identification necessary to provide quality care and the integration of multiple methodologies enabled flexible solutions. District mentors grounded QI coaching and built trust with providers through clinical observation and support.16 A qualitative study of participating QI teams found that the nontraditional collaborative components—equipment and clinical support and the combined mentorship approach—were factors related to high impact QI initiatives.16 Many improvement programs focus on narrow interventions, without taking a comprehensive view of the mortality drivers of MNH care quality across a health system.3,7 Consistent with other research, the Better Birth Trial—a study of the World Health Organization Safe Childbirth Checklist—found that despite improved care practices at individual facilities, there was no impact on mortality, concluding that a greater system improvement focus could be required.23,25

**Variation in Changes Between Districts**

We saw district-level variability in the baseline and change associated with the intervention in some quality measures, consistent with prior study in the region.15 SK had more facilities for the population than Kirehe, with a better developed road system to facilitate access. These factors could have contributed to lower baseline levels in Kirehe for service utilization indicators. We saw rapid change in ANC in Kirehe, which had low documented coverage at baseline and greater room for improvement. However, although Kirehe had a lower mentorship dose due to higher numbers of facilities for the single mentor than in SK, the improvement seen in Kirehe was evidence that ABC could still be associated with meaningful improvement in different contexts.

For immediate skin-to-skin after birth, the baselines were reportedly starkly different between districts. Although this indicator existed in the HMIS before ABC, the register data source was not consistently available at baseline in Kirehe District, which may have contributed to the lower documentation of the practice at the start.

Management of preterm labor with antenatal corticosteroids was a new initiative and began from a low baseline in both district hospitals. Much of the initial work focused on district leadership advocacy to enable dexamethasone procurement by health centers, combined with clinical mentorship to align practice with emerging evidence on safety of steroid administration.26 Therefore, mentors guided clinical providers to provide steroids only when gestational age was known by last menstrual period, which was consistent with national protocol.

Surprisingly, we saw a significant increase in stillbirths in SK in the mixed-effects model. It is possible that despite improvements in target indicators related to improved labor management, stillbirths increased in the face of reduced neonatal deaths if access to quality postnatal care and complications management drove mortality reduction, and quality gaps in antenatal or labor management not captured in our performance measures remained. This possibility is the subject of further investigation.

Despite improved MNH quality of care and neonatal mortality, we found limited change in patient satisfaction. Mutaganzwa et al.19 described the baseline results in detail and found that patient-centeredness of care (including interpersonal relationships, respect, and privacy) and organizational factors such as cleanliness, comfort, and equipment/commodity availability were associated with higher satisfaction with care. Similar to other studies, one explanation for the lack of increased satisfaction corresponding with improved...
quality of care and outcomes could be that women were reluctant to share negative views or had low expectations of facility care and were satisfied with poor quality at baseline. \(^{27}\) It is also possible that the intervention did not have adequate focus directly on improving dignity and respect during patient contact care in the intervention, indicating that future programs may need intentional focus on patient-centered care as part of holistic quality improvement.

Finally, despite the district-level variation in process level impacts, we found reductions in hospital case fatality and district neonatal mortality in the intervention area. The highest impact was seen among preterm/LBW infants, which has been a difficult area for improvement globally. \(^{7,28,29}\) The district-wide approach included supporting specialized care at the district hospital, which likely facilitated this finding and perhaps distinguishing the current study from other studies. \(^{7,30}\) Importantly, Rwanda has experienced health sector improvements over the past decade, \(^{8,31}\) and would be expected to see some reduction in neonatal mortality; however, analysis of available data from the rest of rural Rwanda did not find a similar reduction.

The ABC approach was designed in collaboration with the MOH national and district leadership to enhance scalability and sustainability, which led to the integration of the improvement collaborative components into existing district structures. Based on the results of this program, MOH has continued scaling the approach into non-PIH-supported districts with funding from a Saving Lives at Birth award, with an impact evaluation underway. Furthermore, the program design worked to increase sustainability by building capability of health system actors at multiple levels—district leadership to oversee the approach and integration into routine district systems, and district supervisors to incorporate the mentorship and improvement methodology into their routine work. A sustainability study has been completed to understand system performance 1 year after the end of the intensive period and results will be published separately. As with all health system interventions, integration of effective approaches into routine administrative budgets can be a challenge. Demonstration of the technical impact and program ownership by local leaders were important to support incorporation of the core components (supervisor transport, review meeting costs, clinical and QI continuous learning) into district budgets, and they are ongoing efforts in financially constrained systems.

**Limitations**

Our study had some limitations. Using HMIS and programmatic data meant that we were limited by the quality of the data available. We accounted for this by register review for all possible measures; however, this step would not address issues of accuracy or availability of paper registers. Facilities needed variable lead time to achieve consistent documentation of newly introduced indicators in preterm delivery management, and such documentation was sometimes incomplete. That said, given the growing global call to measure preterm labor management indicators, we hope this practical experience will have global relevance. \(^{32}\)

In addition, mentors coached QI teams to document change ideas and proximal process data rigorously to determine “success” of a given change. However, we cannot be certain that all successful changes were included in the published change package to fully explain the causal pathway to measured improvement in quality of care and mortality. \(^{16}\) Some other time-intensive activities were difficult to capture in program documentation. These included documenting the specific focus of clinical mentorship activities at facilities and the work targeting coordination and communication within and across facilities to strengthen complications management.

Importantly, we did not have a comparison area with measurement and register comparison of mortality; therefore, we used the national HMIS data for rural districts. It is also possible that unknown differences were present between the intervention area and the rest of rural Rwanda. The intervention baseline mortality appears starkly higher than the mortality in the comparison districts, which could be easier to reduce. However, sensitivity analyses using HMIS routinely reported data from the intervention area showed a more comparable baseline mortality, and based on 2010 DHS analyses, the neonatal mortality rates in SK and Kirehe were similar to the rest of rural Rwanda. \(^{12}\) We believe the baseline mortality reported in the comparison districts to be low at least in part due to consistent underreporting in the HMIS, which was overcome in the intervention districts by extracting data from the register records.

Furthermore, improving data quality was part of the QI process, so the difference in differences results could be confounded by change in data quality over time. \(^{33}\) Routine review of HMIS mortality data compared with registers demonstrated underreporting of poor outcomes at baseline,
which decreased throughout the intervention period. If a similar improvement in data quality occurred in the comparison area, it could have reduced the effect size. However, to our knowledge, no concurrent major data quality initiatives were occurring elsewhere in the country.

Mentors triangulated data across registers to use the highest quality data possible for QI and to improve the HMIS data quality. For example, birth weight and gestational age were recorded in the birth register. Mentors compared these data with the counts listed in the HMIS report tally for the same data elements to provide feedback on HMIS data quality to the facility data officer and to build awareness and data quality capability among facility staff in the process. HMIS data QI would not influence the reported quality measures; however, if reporting in the paper registers improved, it would likely lead to underestimation of effect size because it would bias results towards the null in most cases.

Finally, with the introduction of 3 new maternity units in Kirehe District, it is possible that the opening of the 3 new health centers has allowed a decrease of number of patients seeking care in other HCs, which could have contributed to improve the quality of care in the facilities surrounding them by allowing more time per patient; however, we did not track the differences in patient flow.

**CONCLUSION**

We found that the ABC initiative was a feasible and effective program to improve MNH quality of care and reduce neonatal mortality in the intervention districts after 18 months. The QI approach enabled joint problem solving across program and MOH leadership when challenges were encountered. Full transition and further evaluation of sustainability is underway and the Rwanda MOH is currently scaling up ABC into additional non-PHI-supported districts. A mixed-method sustainability analysis will be reported elsewhere. To our knowledge, this study is the first to trace a multi-level neonatal QI program from implementation to clinical process, to mortality impact.24 As countries strive to achieve quality universal health coverage, the ABC initiative could be an important tool for leaders and implementers in countries looking to improve quality and reduce neonatal mortality.

Acknowledgments: We thank the Rwanda Ministry of Health for its leadership in the conception, design, and execution of this initiative and its ongoing support of the scale-up of the quality improvement approach.

Peter Drabac for his leadership and support of the implementation, and the PIH/IMB Monitoring and Evaluation team for their programmatic data collection support. Most importantly, we are grateful to the district leadership, district hospital supervisors, health center heads, and hospital and health center clinical staff, data managers, and community health supervisors for their ongoing commitment to improving the quality of care for mothers and newborns in their communities.

**Funding:** This work was supported with funding from the Doris Duke Charitable Foundation’s Africa Health Initiative (grant no. 2009P001941) and Partners In Health.

**Competing interests:** None declared.

**REFERENCES**


29. Lawn JE, Blencowe H, Ozga S, et al. Every Newborn: progress, priorities, and potential beyond survival. Lancet. 2014;384(9938):189–205. CrossRef. Medline


Peer Reviewed

Received: January 15, 2020; Accepted: August 5, 2020


© Magge et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are properly cited. To view a copy of the license, visit http://creativecommons.org/licenses/by/4.0/. When linking to this article, please use the following permanent link: https://doi.org/10.9745/GHSP-D-20-00031
Where Do Caregivers Take Their Sick Children for Care? An Analysis of Care Seeking and Equity in 24 USAID Priority Countries

Sarah E.K. Bradley, a Lauren Rosapep, a Tess Shiras a

Key Findings

- Out-of-home care-seeking levels were similar for the 3 illness classifications, ranging from 63% for diarrhea to 70% for acute respiratory infection symptoms.
- Among caregivers who sought out-of-home treatment or advice for their sick children, 51% went to the public sector and 43% went to private sector sources.
- There are substantial disparities in illness prevalence and care seeking: 74% of caregivers from the wealthiest households and 63% of caregivers from the poorest households sought out-of-home care for their sick children.

Key Implication

- Stewards of the public and private sectors—including governments, nongovernmental organizations, civil society, donors, and implementing partners—must understand sources of sick child care and collaborate to improve access to and quality of affordable care.

ABSTRACT

Pneumonia, diarrhea, and malaria are leading causes of under-5 mortality. Accelerated reductions in illness burden are needed to meet childhood Sustainable Development Goals. Understanding where parents take sick children for care is key to improving equitable, high-quality treatment for these childhood illnesses and catalyzing reductions in morbidity and mortality. We analyzed the most recent Demographic and Health Survey data in 24 of the United States Agency for International Development’s maternal and child health priority countries to examine levels and sources of care for children sick with 3 illness classifications: symptoms of acute respiratory infection, diarrhea, or fever. On average, across countries analyzed, one-third of children had recent experience with at least 1 of the 3 classifications. The majority (68.2%) of caregivers sought external advice or treatment for their sick children, though the level is far higher for the wealthiest (74.3%) than poorest (63.1%) families. Among those who sought out-of-home care, 51.1% used public sources and 42.5% used private-sector sources. Although sources for sick child care varied substantially by region and country, they were consistent across the 3 illness classifications. Urban and wealthier families reported more use of private sources compared with rural and poorer families. Though 35.2% of the poorest families used private sources, most of these (57.2%) were retail outlets like pharmacies and shops, while most wealthier families who sought care in the private sector went to health facilities (62.4%). Efforts to strengthen the quality of integrated management of sick child care must therefore reach both public and private facilities as well as private pharmacies, shops, and other retail outlets. Stakeholders across sectors must collaborate to reach all population groups with high-quality child health services and reduce disparities in care-seeking behaviors. Such cross-sectoral efforts will build clinical and institutional capacity and more efficiently allocate resources, ultimately resulting in stronger, more resilient health systems.

INTRODUCTION

In the last 30 years, there has been remarkable global progress in reducing child mortality. The under-5 mortality rate decreased by more than half, from 93 deaths per 1,000 live births in 1990 to 39 deaths per 1,000 live births in 2018. Yet, tremendous work remains. On average, 15,000 children died every single day in 2018.
Progress in reducing child mortality has also been uneven: in an analysis across 137 low- and middle-income countries (LMICs), the under-5 mortality rate was more than 2 times higher among children from the poorest than the wealthiest families (65 versus 31 deaths per 1,000 live births, respectively). To meet global child health goals, the development community will need to focus on closing equity gaps and accelerating further reductions in child deaths, including those from pneumonia, diarrhea, and malaria. Together these 3 preventable and treatable illnesses account for nearly one-third of under-5 deaths.

Many LMIC governments and donors are focusing on countries’ journeys to self-reliance to simultaneously transition away from donor dependence and increase gains in maternal and child survival. Key to this goal is collaboration and harmonization of efforts across all health sector actors. Global development stakeholders including the Gates Foundation, the United States Agency for International Development (USAID), United Kingdom’s Department for International Development, and the Global Financing Facility have emphasized the importance of collaborating with governments, other donors, civil society, faith-based organizations, and the private sector effectively and efficiently to save women and children’s lives. USAID’s Private Sector Engagement Policy echoes and amplifies this message, calling on public and private sector actors to:

take the unique capabilities of each [sector] and apply them to problems that neither could address fully on their own.

To facilitate such collaboration and harmonization, stakeholders first need to understand whether and where parents are seeking treatment for their sick children and how levels and sources vary across regions, countries, and population groups.

This study aims to provide updated information on levels and sources of care for children sick with symptoms of acute respiratory infection (ARI), diarrhea, and/or fever with an equity lens. We analyzed data from the most recent Demographic and Health Survey (DHS) data for 24 of the 25 USAID maternal and child health priority countries to address 3 research questions:

1. What is the prevalence of reported ARI symptoms, diarrhea, and/or fever among children under 5 in USAID priority countries?
2. How commonly do caregivers seek out-of-home care for their sick children?
3. When caregivers seek treatment or advice for their sick children, which sources do they use?

We examined equity implications for each research question and present notable differences in prevalence rates and care-seeking patterns by urban-rural residence and between households in the highest and lowest wealth quintiles of each country. We recognize that there is a great degree of overlap between socioeconomic status and urban-rural residence. For example, in 21 of the 24 priority countries analyzed, more than 90% of the poorest families live in rural areas. However, barriers related to these characteristics are different: care-seeking barriers related to geography are more likely to pertain to availability of and access to services whereas barriers related to socioeconomic status are more likely related to affordability. Therefore, we disaggregated findings by both residence and socioeconomic status to allow practitioners to understand differences in these populations and draw implications for programs and policies.

In the context of an increasing focus on private-sector engagement in health and using newly-available Demographic and Health Survey (DHS) data, we have built on previous analyses to provide the most up-to-date information on levels and sources of sick child care. We examined patterns across population segments with a focus on socioeconomic equity, noting the substantial progress that needs to be made. At the end of this paper, we describe and provide links to additional interactive data visualization resources and describe some of our work to date promoting the use of this information to transform research into action at the country level.

■ DATA AND METHODS

Data

We analyzed the most recent nationally representative DHS household survey data from each priority country that was available on the DHS website as of December 31, 2019 (Figure 1). Data were available for all priority countries except South Sudan. Latest available data are quite recent for most countries except for Madagascar (from 2009), Mozambique (from 2011), and Yemen (from 2013). For these 3 countries, we note that data presented here may not accurately reflect the current situation.

Data on Prevalence of Illness Classifications

During DHS data collection, eligible women are interviewed and asked about each of their children. For each living child aged 5 years or
younger, mothers are asked questions about whether their child had cough, diarrhea, or fever in the 2 weeks before the survey. If the child had a cough, the mother is asked whether the child had rapid or difficult breathing that was chest-related. Following DHS standards, we classify cough with rapid or difficult chest-related breathing as ARI symptoms, which are used as a non-specific proxy for pneumonia. Fever is used as a non-specific proxy for malaria. In this article, we refer to ARI symptoms, diarrhea, and fever as illness classifications, which are not confirmed by any diagnostic tests, but come directly from mothers’ reports.

Data on Care Seeking and Sources for Care
If women report that 1 or more of their children was ill in the last 2 weeks, they are asked if they sought treatment or advice for each child from any source. Care seeking is classified as any care sought outside the home. Because we focused on sources of care, this analysis focuses on care outside the home, not whether the child received an appropriate treatment (e.g., oral rehydration solution and zinc supplements, which could have been administered at home). Similarly, this analysis likely does not capture other medications kept and administered at home, whether or not they are appropriate.

If mothers report seeking treatment or advice outside the home, they are asked where they went. Their responses are classified into precoded categories that vary by country. To standardize categories across countries, we classified sources into public or private sector or “other” sources including traditional healers, friends, and family (Table). Public and private sector sources are further classified by whether the source was a health facility, a community health worker (public sector only), or a retail outlet (private sector only). We note that these categories, based on mother’s recall, may not perfectly capture precise sources, but we believe the categories used are broad enough to represent source groups with reasonable accuracy. We also note that these results reflect where sick child care was sought in the 2 weeks before the survey and may not reflect parents’ subsequent care-seeking destinations or where parents may prefer to seek care if barriers (geographic, financial, etc.) were removed.
Data on Equity
We examined results by socioeconomic status by using the DHS wealth quintiles, which divide the population surveyed in each country into evenly-sized quintiles based on their household assets. \(^1\) We used the bottom and top quintiles, respectively, to represent children and caregivers from the poorest 20% and wealthiest 20% of households in each country. We also examined results by urban and rural residence. We used the DHS classifications of urban and rural, which are based on the classifications used in each country.

Analytic Methods
The unit of analysis for all research questions is children aged 0–59 months (under 5 years) old. All analyses used DHS survey sampling weights. To generalize results across countries, we multiplied the survey weights by a survey-specific constant to standardize the effective weighted sample size across countries. Thus, each country contributes equally to the regional and all-country averages, and results are not weighted more heavily toward surveys with larger sample sizes or populations. We considered weighting results by the population size of each country, but found that nearly three-quarters of the population-weighted sample would be from Asia because the Asian countries in our analysis are more populous, and noted that the 24 countries are not representative of any regions or larger geographies. Therefore, average estimates should be interpreted as averages across countries analyzed. Similarly, regional results are not representative of the entire region but should be interpreted as the average across countries analyzed in each region. All surveys are included in averages, but country-level results are suppressed if they are based on fewer than 50 unweighted cases. All analyses and visualizations were conducted in Stata version 14.2.

RESULTS
Results presented here describe the prevalence of ARI symptoms, diarrhea, and/or fever among children 5 or younger and associated care-seeking levels and source patterns. In each section, we present regional averages for countries analyzed in Asia, East and Southern Africa, and West and Central Africa to summarize patterns observed in the data.

Prevalence of Illness Classifications
Reports of ARI symptoms, diarrhea, and fever among children under 5 were common across the countries examined. The prevalence of these illness classifications ranged widely between countries (15.9%–46.9%). On average, 1 of 3 children (32.9%) experienced 1 or more of these 3 illness classifications in the 2 weeks before the survey.

Fever was the most common classification reported in all priority countries and regions. On average across priority countries, 23.4% of children under 5 experienced fever, 15.5% had diarrhea, and 5.8% experienced ARI symptoms (Figure 2).

Across all countries analyzed, the prevalence of fever (Figure 2) ranged widely: Madagascar had the lowest prevalence of fever (9.3%); both Bangladesh (36.8%) and Pakistan (37.6%) had the highest. Bangladesh was an outlier with a comparatively high prevalence of fever and for its low level (5.6%) of diarrhea. Yemen (31.2%) and Afghanistan (28.7%) stand out for diarrhea prevalence that was double the all-country average (15%). With the lowest levels on average, the range of ARI prevalence across countries was
not as large; Nepal (2.4%), Mali (2.0%), and Mozambique (1.5%) were the countries with the lowest ARI prevalence, and Pakistan had the highest rate (13.8%).

Prevalence of Comorbidities

Comorbidities—reports of multiple types of illness at the same time—were common among children in priority countries. Among the 3 classifications examined, nearly 1 in 5 (18.8%) children suffered from more than 1 symptom in the 2 weeks before the survey. Regionally, comorbidities were highest in the Asian countries analyzed (21.4%), compared to 15.9% in the sub-Saharan countries. Children in Haiti (34.9%) and Pakistan (34.1%) had the highest prevalence of comorbidity.

Equity Implications: Disparities in Illness Classification Prevalence

Differences in illness classification prevalence between children living in urban and rural areas was small (less than 4 percentage points for all regions) but were more notable between children from the poorest and wealthiest households. On average, across all countries and all 3 illnesses classifications, prevalence was 4.5 percentage points higher among children in the poorest households (34.1%) than those in the wealthiest households (29.6%). By region, the disparity in prevalence of 1 or more illness classifications was largest in West and Central African countries (8.7 percentage points), followed by East and Southern Africa (3.9 points), and smallest in the countries analyzed in Asia (2.1 points), with substantial variation at the country level. The disparity for children in the poorest and wealthiest households in most countries was 10 percentage points or lower; the notable outliers were Nigeria and Uganda where the disparity between the wealthiest and poorest (22.9 and 20.7 percentage points, respectively) was more than double the disparity of any other country examined (Figure 3).

The magnitude of the disparity in illness classification prevalence between children from the poorest and wealthiest households differed somewhat by type of classification. The largest disparities were observed for fever: in West and Central African countries analyzed fever prevalence was 6.4 percentage points higher on average among the poorest households compared to the wealthiest households, and 4.8 percentage points higher among poorest than wealthiest households on average in the East and Southern African countries analyzed. Disparities in diarrhea prevalence were generally smaller than those for fever. The largest reported diarrhea prevalence disparities were observed in the West and Central African countries analyzed, particularly Nigeria and Senegal (12.5 and 9 percentage points, respectively). Outside of this region the average disparity for reported diarrhea prevalence was 1.4 percentage.
points in the East and Central African countries analyzed and was 1.2 percentage points in the Asian countries analyzed. The smallest disparities between poorest and wealthiest children observed in priority countries were for reported ARI symptoms: 1.7 percentage points on average across all countries analyzed, with West and Central African countries having the smallest average disparity (0.9 percentage points). For more details, see Supplemental Figures.

Out-of-Home Care Seeking

On average across all USAID priority countries, most caregivers (68.2%) sought treatment outside of the home when their children experienced 1 or more of the 3 classifications examined in this analysis. Overall, care seeking was highest in Asia (77.8%) with lower levels across Africa (66.9% in West and Central Africa, 64.5% in East and Southern Africa). At the country level, care-seeking...
levels varied widely (Figure 4) with the lowest levels of care seeking in Ethiopia (38.8%) and Madagascar (44.2%) and the highest levels of care seeking in Indonesia (88.9%) and Bangladesh (83.6%).

**Care Seeking by Illness Classification Type**

Although care-seeking levels vary substantially across countries, they are similar across illness classifications. Care-seeking levels were highest for ARI (70.4%) and fever (68.5%) and slightly lower for diarrhea at 63.3%. This pattern, which was also observed at the regional level, may be partially driven by the fact that diarrhea can be effectively managed at home with oral rehydration solutions and zinc supplements.

**Equity Implications: Gaps in Care Seeking by Urbanicity and Wealth**

Across all countries analyzed there were equity-related gaps in care seeking. The care-seeking level in rural areas was 6.3 percentage points lower on average than the urban care-seeking level, with a high-degree of variability across countries (Figure 5). In general, differences were smallest where overall care-seeking levels were highest. For example, in Indonesia, 89% of both urban and rural caregivers sought sick child care outside the home. The largest urban-rural care-seeking gaps were in countries with lower overall levels of care seeking, particularly in Madagascar where care was sought for 41.2% of rural children and 61% of urban

---

**FIGURE 4.** Out-of-Home Care-Seeking Levels by Country Among Caregivers of Sick Children in USAID Maternal and Child Health Priority Countries

<table>
<thead>
<tr>
<th>Country</th>
<th>Care-Seeking Level (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indonesia 2017</td>
<td>88.9</td>
</tr>
<tr>
<td>Bangladesh 2014</td>
<td>83.6</td>
</tr>
<tr>
<td>India 2015-16</td>
<td>80.6</td>
</tr>
<tr>
<td>Uganda 2016</td>
<td>80.5</td>
</tr>
<tr>
<td>Pakistan 2017-18</td>
<td>79.4</td>
</tr>
<tr>
<td>Tanzania 2015-16</td>
<td>78.7</td>
</tr>
<tr>
<td>Nepal 2016</td>
<td>77.8</td>
</tr>
<tr>
<td>Liberia 2013</td>
<td>77.6</td>
</tr>
<tr>
<td>Ghana 2014</td>
<td>76.1</td>
</tr>
<tr>
<td>Nigeria 2018</td>
<td>73.9</td>
</tr>
<tr>
<td>Zambia 2013-14</td>
<td>73.3</td>
</tr>
<tr>
<td>Kenya 2014</td>
<td>72.1</td>
</tr>
<tr>
<td>Malawi 2015-16</td>
<td>70.4</td>
</tr>
<tr>
<td>Myanmar 2015-16</td>
<td>69.2</td>
</tr>
<tr>
<td>Afghanistan 2015</td>
<td>65.2</td>
</tr>
<tr>
<td>Yemen 2013</td>
<td>63.7</td>
</tr>
<tr>
<td>Mozambique 2011</td>
<td>62.6</td>
</tr>
<tr>
<td>DRC 2013-14</td>
<td>61.0</td>
</tr>
<tr>
<td>Rwanda 2015-16</td>
<td>59.8</td>
</tr>
<tr>
<td>Mali 2018</td>
<td>60.4</td>
</tr>
<tr>
<td>Senegal 2017</td>
<td>52.6</td>
</tr>
<tr>
<td>Haiti 2016-17</td>
<td>46.7</td>
</tr>
<tr>
<td>Madagascar 2009</td>
<td>44.2</td>
</tr>
<tr>
<td>Ethiopia 2009</td>
<td>38.8</td>
</tr>
</tbody>
</table>

Abbreviations: DRC, Democratic Republic of Congo; USAID, U.S. Agency for International Development.
children—a gap of 19.8 percentage points—and in Ethiopia, which had the lowest overall levels of care seeking and the greatest urban-rural disparity, at 26.5 percentage points.

There were also substantial disparities in care seeking for children from the poorest and wealthiest households (Figure 6). Across all countries, the average care-seeking level for any illness classification was 11.3 percentage points higher for caregivers from the wealthiest households (74.4%) than from the poorest households (63.1%). Among the Asian countries analyzed, the disparity in care seeking ranged from 3.0 to 11.0 percentage points. Sub-Saharan African countries, on the other hand, had a much wider range in care-seeking gaps. In some cases (Zambia, Kenya, Ghana, and Malawi) care seeking among the poorest caregivers was similar to or higher than care seeking among the wealthiest caregivers, but in other cases (Rwanda, Madagascar, Senegal,
Ethiopia) the wealth gap in care seeking exceeded 22 percentage points or double the size of the gap for priority countries overall.

**Sources of Care**

On average, across countries analyzed, half of caregivers (51.1%) who sought out-of-home care for their children’s most recent illness reported that they went to a public source; 42.5% reported that they went to a private sector source, 4.9% reported that they sought care from “other” sources (including informal providers, traditional healers, friends, or family members), and 1.5% reported that they sought care from both public and private sources. This care-sourcing pattern was very consistent across classifications; for example, the proportion of caregivers who used private sector sources was 41.2% when their child had diarrhea, 43.6% for fever, and 42.2% for ARI symptoms. This consistency in both levels of care seeking and treatment sources

---

**FIGURE 6. Differences in Care Seeking Among Caregivers With Sick Children From the Poorest and Wealthiest Households in USAID Maternal and Child Health Priority Countries, by Country**

The bars depict the magnitude of the difference between the poorest and wealthiest, with values shown in cases where the magnitude is 5 or more percentage points.
by illness classification may indicate that the symptoms children were experiencing did not factor heavily in caregivers’ decision making about where to seek out-of-home advice and treatment. Because sources did not vary by classification, the remaining results in this section focus on sources used for children sick with any of the 3 illnesses classifications.

However, sources of sick child care did vary within and across regions (Figure 7). In Asian countries analyzed, the private sector was the dominant source of care, with 59.9% of caregivers seeking care from this source. Pakistan had the highest level of private sector care seeking (80.4%) of any country analyzed, and Myanmar and Afghanistan’s private sector care-seeking levels (39.3% and 38.2%, respectively) were substantially lower than other countries in that region. By contrast, in East and Southern Africa, the public sector was the source of care for a large majority (69.6%) of sick children, and in Mozambique, the public sector was the source of care for almost all (91.9%) children. In West and Central Africa, there was a more even split in care seeking from public (52.1%) and private (42.5%) sources on average. There was much variation at the country level, for example, in Nigeria where the private sector was the source of out-of-home care for the majority (60%) of caregivers. Although relatively few caregivers sought care from other informal sources, there were 2 notable exceptions: in Bangladesh, the level of care sought from other sources (predominantly “unqualified doctors”) was 30.5%, and in Mali it was 13.8% (predominantly “traditional practitioners”).

**FIGURE 7.** Sources for Sick Child Care in USAID Maternal and Child Health Priority Countries, by Country

Abbreviations: DRC, Democratic Republic of the Congo; USAID, U.S. Agency for International Development.
Equity Implications: Comparing Sources of Care by Urbanicity and Wealth

Analysis of reported sources for sick child care across all countries and regions showed similar patterns when disaggregated by urban versus rural residence (Figure 8). In each region, caregivers in rural areas had lower levels of care seeking from private sources and higher levels of care seeking from public and informal sources.

The patterns of sources of care used by caregivers from the poorest and wealthiest quintiles was nearly identical to sources used by rural and urban caregivers: a consistent pattern of higher private sector use among the wealthiest and higher public and informal sector use among the poorest caregivers in each region. Despite this overall pattern, analysis across countries revealed that caregivers of all wealth levels— even the poorest caregivers—sought care for their children from private sector sources. On average across priority countries, 35.4% of caregivers from the poorest households in each country sought care from private sector sources and 55.1% of caregivers from the wealthiest households.

Types of Public and Private Sector Sources

The public and private sectors are not homogenous and are made up of different types of health providers. In the public sector, sources included health facilities (hospitals, clinics, health posts) and community health workers (CHWs). In the private sector, reported sources included health facilities (hospitals, doctors, and private for-profit, nongovernmental organizations, and faith-based clinics) and private retail outlets (pharmacies, shops, and markets). When seeking care from a health facility, caregivers are likely to interact with clinically trained health professionals. In contrast, prior research has shown that retail outlets may be less likely to have adequately trained health professionals, and providers may have limited access to or training on current treatment and counseling policies and guidelines, which could potentially result in a substandard quality of medical care.13–15

Among caregivers who recently sought care from public sector sources, nearly all (95.6%) reported that they received care at a health facility rather than from a CHW. Fewer than 5% of caregivers in most countries reported that they sought help from a CHW. However, there were 2 notable outliers; in Indonesia and Rwanda, 26.7% and 26.3% of caregivers, respectively, sought care or advice from a CHW.

Overall, care seekers who used the private sector consulted with providers in health facilities (48.8%) and retail outlets (52.4%) at nearly equal levels. Although this split did not vary by illness classification, there was somewhat more variation at the regional level. In the Asian and East and

FIGURE 8. Sources for Sick Child Care in USAID Maternal and Child Health Priority Countries, by Region and Urbanicity

Abbreviations: DRC, Democratic Republic of the Congo; ESA, East and Southern Africa; USAID, U.S. Agency for International Development; WCA, West and Central Africa.
Southern African countries analyzed, a majority of caregivers who reported private-sector care sourcing went to a health facility (61.4% and 55.7%, respectively). By contrast in the West and Central African countries analyzed, a minority of caregivers (22.1%) reporting private sector care seeking went to a health facility.

**Equity Implications: Comparing Private Sources of Care by Urbanicity and Wealth**

Among caregivers who sought care from private sector sources, the types of providers seen varies by urban/rural residence and socioeconomic status. This section examines the use of health facilities versus pharmacies and retail outlets among caregivers who used private sector sources. By urban and rural residence, health facility use versus retail outlet use followed a similar pattern across all regions (Figure 9). In each region, caregivers residing in rural areas had lower levels of care seeking from private health facilities and higher levels of care from private retail outlets. This pattern may reflect a distribution of health facilities that favors urban areas.

Similarly, we found that within the 35.4% of the poorest caregivers and 55.1% of the wealthiest caregivers who sought care from the private sector, the poorest and wealthiest used different sources within the sector. On average across countries, 62.4% of the wealthiest caregivers who reported private sector care seeking went to a health facility. The data reflect a converse scenario for those with the lowest socioeconomic status; among this segment a majority (57.2%) of the poorest private sector care seekers used retail outlets. On average, the gap in private health facility care seeking between the wealthiest and poorest caregivers was 33.5 percentage points in the Asian countries analyzed and 31 percentage points in the West and Central African countries analyzed. The average gap in private health facility care seeking in the East and Southern African countries analyzed was smaller (17.9 percentage points). Except for Haiti, this overall pattern holds for all priority countries (Figure 10). Disparities were particularly large in Madagascar and Bangladesh where the gap in private health facility care seeking between the wealthiest and the poorest households was 51.4 and 42.2 percentage points, respectively.

**DISCUSSION**

Reports of treatable and preventable illness symptoms were extremely common across USAID’s maternal and child priority countries, affecting 1 in every 3 children under age 5. Given this high illness burden, stewards of the public and private sectors—including nongovernmental organizations, civil society, donors, and implementing institutions—must consider the implications for their health systems and service delivery strategies.
Both public and private sectors were critically important for families seeking care, reinforcing the need for sectors to complement each other to provide equitable access to high-quality care.

Although the majority of caregivers (68.2% on average) sought treatment or advice when their children were sick, there are still many areas where progress is needed, especially to improve equity. For example, in nearly half of the priority countries, more than one-third of the poorest caregivers did not seek care outside the home for their sick children. On average across countries, there was a 6.3 percentage point disparity in care-seeking levels comparing rural with urban families and an 11.3 percentage point disparity between the wealthiest versus poorest families. Potential barriers to seeking timely advice and treatment may include poor availability, limited access, unaffordability, and inadequate information about illness danger signs. In countries with overall low care seeking or low care seeking among specific...
population groups, additional exploration into these potential barriers is warranted.

Among those families that did seek care, both the public (51.1%) and private (42.5%) sectors were critically important, reinforcing the need for health actors across sectors to complement one another and ensure that there is equitable access to high-quality care for all population segments. While the majority of the poorest (56.6%) and rural (52%) care seekers accessed advice and treatment from public sources, more than 1 in 3 of the poorest (35.4%) and rural (37%) care seekers used private health facilities, pharmacies, or shops. Poorer and rural families may rely on private sources due to increased convenience, quality perception, or even affordability, for example if subsidized products are available in the private nonprofit sector.

Although use of private retail outlets like pharmacies and shops was quite high (52.3%) overall among care seekers who used the private sector, the poorest care seekers in this group sought much more care (59.2%) from retail outlets than wealthy care seekers who used the private sector (37.4%). This is an important disparity because these types of retail outlets may be less likely to be staffed by providers trained with up-to-date guidelines, suggesting that the poorest care seekers may be at risk of obtaining poorer quality care than their wealthier counterparts. As efforts expand in the private sector to strengthen the quality of integrated management of sick child care, this finding is key to ensure that these efforts reach both private facilities as well as private pharmacies, shops, and other retail outlets.

Among care seekers who sought care in the public sector, the large majority (95.6%) consulted health facilities rather than CHWs. This aligns with results from a Hodgins et al. analysis (2013) in which they concluded that it is inappropriate to focus program efforts on community health workers to the exclusion of more widely used sources of care.

When sharing our analysis results with stakeholders in a selection of priority countries, stakeholders were particularly surprised to see low use of CHWs given the amount of resources dedicated to community case management. A potential limitation pertaining to reporting of CHWs is that DHS asks women where they “go,” so respondents may not report a CHW if they saw a CHW at a health post or if they did not physically “go” outside of their home for care. In addition, it is possible that CHWs provide advice but do not consistently stock appropriate treatment, so caregivers may go where they know supplies can be readily obtained. Before refocusing program efforts away from CHW case management programs, we suggest conducting country-specific explorations into CHW programs to determine if and how they are used and barriers to successful operation.

Using the Data

It is critical to examine childhood illness classifications and care-seeking patterns at the country level where policy makers, advocates, and civil society are charged with implementing cost-effective and sustainable policies and programs that will lower childhood mortality and meet Sustainable Development Goals. These data are pertinent for government actors employing strategies to efficiently mobilize domestic resources across sectors. Interpreting findings at the country level requires a robust understanding of national and subnational factors including health financing, health system functioning, governance and policy, sociocultural norms, and ongoing campaigns or interventions. With this contextual knowledge, stakeholders can use these data to better understand and begin to tackle challenges such as low care-seeking levels, inequities in care seeking, and poor access to particular sources of care.

To assist country stakeholders in interpreting and using these findings, we have presented data through accessible infographics and data visualizations. All of our data are housed in an online interactive data visualization tool—called PrivateSectorCounts.org—that allows users to explore findings in their own country or across countries and make comparisons across demographic characteristics. Further, we have disseminated findings through country-specific briefs and annotated PowerPoint presentations, available on SHOPPlusProject.org.

The authors worked with several USAID Missions to interpret findings together based on their nuanced country knowledge and subsequently disseminated analysis findings together via 3 regional webinars. This process yielded fruitful insights and programmatic recommendations, exemplified through the quotes below:

*One key thing . . . is advocating for expanding health financing options to deliver services. These results show that the private sector has a big role in delivering services, and the government needs to put in more effort to ensure that they strengthen the private [sector] equally to the public sector. . . . These results give us a powerful tool to advocate for a comprehensive health financing strategy.* —USAID/Uganda Representative
In India, it’s surprising that even 70% of the poorest seek care from the private sector ... Our understanding was that amongst the poor, the public sector was dominant. But, this data shows us that poor also get care from the private sector. This gives us a clue about improving the care in the private sector. —USAID/India Representative

**Limitations**

DHS data are self-reported, so it is possible that there are errors or misclassifications in sources and levels of care used. For example, it is possible that respondents do not consider interactions such as buying medication in a shop to be “advice or treatment outside the home,” which could lead to underreporting of care-seeking levels, though given the high levels of reported care seeking from pharmacies and retail outlets, we anticipate any underreporting would likely be minimal. In addition, respondents are asked where they “go” for care, not whom they see, so if respondents see a CHW at a health post, for example, the response may be recorded as health post (categorized as a type of health facility) rather than as a CHW. The low levels of reported CHW care provision may also be related to the fact that the DHS only includes a standard response category for public sector CHWs, rather than response categories both for public and private sector CHWs. Local understandings of who is a CHW (some of whom might be reported by respondents as “traveling nurses” instead, for example) may further complicate efforts to understand care-seeking levels through CHWs via this analysis alone. We also note that data were collected from mothers, though a different caregiver may have been the one to seek care—this is especially likely in countries where women often have restricted mobility as in Afghanistan. If a different caregiver sought care for the sick child, the mother may not have had complete information on care-seeking sources.

Additionally, this analysis does not include data about preferences for sources of care. The sources reported are those used and may not be the preferred source of care if all sources were available and accessible to the respondent. As such, conclusions cannot be drawn about preferred care sources from these data.

**CONCLUSION**

The public and private sectors both play key roles in treating sick children. Stakeholders across sectors must collaborate and strategize to reach all population segments with high quality child health services and work toward reducing disparities in care-seeking behaviors. Given the high use of private retail outlets—namely pharmacies, drug shops, and markets—efforts to ensure knowledge of and adherence to appropriate integrated management of childhood illness protocols in these outlets should continue and be strengthened, including through additional research when warranted. Cross-sectoral communication and joint problem solving is particularly critical in the context of the COVID-19 pandemic and other external shocks that create health system and health care-seeking constraints. Such cross-sectoral efforts will build clinical and institutional capacity and more efficiently allocate resources, ultimately resulting in stronger, more resilient health systems.

**Acknowledgments:** The authors gratefully acknowledge the technical guidance received from Catherine Clarence and analytical assistance from Anthony Leegwater, both of Abt Associates, and helpful feedback and review from Malia Baggs and Nefra Faltas of USAID.

**Funding:** This work was funded through Sustaining Health Outcomes through the Private Sector (SHOPs) Plus, a 5-year cooperative agreement (AID-OAA-A-15-00067) funded by the United States Agency for International Development (USAID). The project strategically engages the private sector to improve health outcomes in family planning, HIV, maternal and child health, and other health areas.

**Competing interests:** None declared.

**REFERENCES**


Peer Reviewed

Received: March 18, 2020; Accepted: July 15, 2020; First published online: September 4, 2020


© Bradley et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are properly cited. To view a copy of the license, visit http://creativecommons.org/licenses/by/4.0/. When linking to this article, please use the following permanent link: https://doi.org/10.9745/GHSP-D-20-00115
A Qualitative Comparative Analysis of the Drivers of HIV Status Knowledge in Orphans and Vulnerable Children in Mozambique

Allie Davis, Zola Allen, Nena do Nascimento, Jenifer Chapman, Rotafina Donco, Daan Velthausz

Key Findings

- Using fuzzy-set qualitative comparative analysis, we identified combinations of modifiable factors that HIV programs that support orphans, vulnerable children, and their families may be able to act on to increase the proportion of beneficiaries who know their HIV status.

Key Implications

To improve HIV status knowledge, programs should consider implementing a formal process to:

- Assign cases considering case complexity and caseload to reduce overwork
- Provide activistas with supervision support
- Hire experienced activistas
- Train activistas to address challenging cases
- Reimburse activistas for work-related expenses

Abd

Resumo em português no final do artigo.

ABSTRACT

In Mozambique, more than a million children are living with HIV or are otherwise vulnerable due to HIV. In response to this crisis, the US President’s Emergency Plan for AIDS Relief funds programs that serve orphans and vulnerable children affected by HIV and their families. These programs retain case workers, known as activistas, who provide and refer beneficiaries to services to increase beneficiaries’ knowledge of their HIV status and to improve retention in care among those living with HIV. To improve program effectiveness, implementing organizations need to understand how different case management attributes affect beneficiary outcomes. We applied fuzzy-set qualitative comparative analysis (QCA), based on 119 interviews, to identify the combinations of case management attributes that led to (1) increased knowledge of HIV status, and (2) high percentages of beneficiaries with known HIV status. We identified 6 pathways for the first outcome and 5 pathways for the second outcome. Each pathway demonstrates an alternative combination of conditions that positively influences the outcome and is equally sufficient in achieving the outcome. To improve knowledge of HIV status, programs in a similar context as this study may select any of the identified pathways based on their existing resources and work on ensuring the presence of each of the conditions in the pathway. Overall, based on the presence of some of the factors in multiple pathways, we conclude that to improve knowledge of HIV status it is important that programs implement a formal process to assign cases considering case complexity and existing caseload to reduce overwork; provide activistas with external support such as weekly care team meetings, weekly supervisor meetings, and/or low supervision ratios; hire experienced activistas; provide all activistas with follow-up trainings so they have the tools to address challenging cases; expand the financial resources offered to activistas; and reimburse activistas for work-related expenses.

INTRODUCTION

The HIV epidemic has exacted a formidable toll on children and their families. Currently, 13.4 million children are living without 1 or both parents due to the HIV epidemic; 80% of these children live in sub-Saharan Africa. In addition, 1.8 million children under age 15 are...
living with HIV. Despite some decline in HIV adult prevalence worldwide and increasing access to treatment, the number of children affected by or vulnerable to HIV remains alarmingly high. In response to the global HIV/AIDS epidemic, the United States initiated the President’s Emergency Plan for AIDS Relief (PEPFAR) in 2003. PEPFAR funds health care facilities, nongovernmental organizations, and other programs that provide support and services to populations vulnerable to or infected with HIV. In many countries with a high burden of HIV, PEPFAR supports programming for treatment and support to orphans and vulnerable children (OVC) and their families, especially those made vulnerable by HIV.

The primary mechanism for service delivery and support for OVC programming is client case management. Individual orphans, vulnerable children, or their family members are enrolled into a program, and a case worker is assigned. The case worker assesses client needs, outlines a care plan and actions to achieve that care plan, monitors care plan achievement, and ultimately exits the client from the program upon care completion. A critically important aspect of case management is that case managers advocate on behalf of their clients, ensuring that medication is received, patient calls are returned, paperwork is filed with the clinic or care agency, and HIV- and non-HIV-related needs are met. Case workers are volunteers or stipend-paid community members who are trained by the program to provide services to clients, but they otherwise have little or no formal social work training or qualifications. Case workers are employed by a community-based organization (CBO) or implementing partner. CBOs often employ case worker supervisors whose role is to review client files with case workers, support them in meeting clients’ needs, support time management, assess training needs, identify training opportunities, and provide support to help them cope with job stress.

Client case management has been shown to positively influence key HIV outcomes, notably knowledge of HIV status and antiretroviral therapy (ART) retention. Studies have shown that increased contact with a case manager is related to a decreased need for income assistance, emotional counseling, and other forms of supportive services and increased access to health care. HIV programs are increasingly turning to family-oriented, case management approaches because of their potential to provide holistic, individualized, and effective support, particularly among vulnerable populations.

Although case management is widely accepted as an effective method for HIV service provision, little is known about the specific attributes that make a case management program effective. Most studies that have evaluated case management programs have focused solely on how well these programs achieved final outcomes, such as knowledge of HIV status, ART retention, or cost effectiveness; they did not investigate the factors that led to those outcomes. The limited number of studies that have examined specific case management attributes primarily consider a small number of factors in isolation and do not account for how factors may combine to have a collective impact on program outcomes. From previous studies, several case management program attributes have been hypothesized to positively influence case management effectiveness, such as case manager skills, training, supportive supervision, financial incentives, resources, networking, and accessibility of care. Although there are many factors posited to influence knowledge of HIV status, such as stigma, demographics (e.g., beneficiary age, sex, income, education), and social capital, this study focused solely on the modifiable attributes of case management programs because these factors are in programs’ manageable control. The limited research on case management attributes that improve HIV outcomes highlights the need for research that comprehensively and specifically evaluates which aspects of case management programs influence case management effectiveness.

We used qualitative comparative analysis (QCA) to identify the modifiable case management attributes that optimize program performance within a PEPFAR-funded program in Mozambique. We defined optimized program performance as a scenario in which the proportion of beneficiaries that know their HIV status is increasing, which is a commonly used indicator in HIV programming. Specifically, this study identified the combinations of modifiable case management attributes that influenced 2 indicators (i.e., outcomes) of effectiveness: (1) percentage change in knowledge of HIV status, and (2) percentage of beneficiaries with HIV status known at the time of the last assessment. The results of this study contribute to a theory of effective OVC programs, identifying actionable recommendations that implementing organizations can follow to optimize program performance.

## METHODS

### Research Context and Case Selection

In Mozambique, approximately 15% of women and 10% of men ages 15–49 are living with HIV.
Under the age of 15, approximately 200,000 children in Mozambique are living with HIV, and 916,000 are considered vulnerable because of HIV prevalence. Low rates of treatment retention, especially among children, adolescents, and young adults, threaten to undermine epidemic control. Due to the high prevalence and risk for HIV in Mozambique, the United States Agency for International Development (USAID) funds OVC-focused programming through a program called COVida. The program partners with CBOs to recruit and equip case managers to provide services to OVC and their families and is a major center of case management in Mozambique. COVida-affiliated CBOs are the focus of this study. The unit of analysis for this study is the case manager, called activista in Mozambique (and referred to as such herein). Activistas are expected to work about 20 hours weekly. Across all COVida CBOs, activistas have managed 344,000 beneficiaries, 60% of whom are currently active.

For QCA, it is important that the units of analysis exhibit varied degrees of the outcome(s) and the factors analyzed. COVida has CBOs in all 11 provinces in Mozambique. To gather in-depth knowledge about each program, we selected 3 provinces: Maputo, Gaza, and Nampula. The 3 provinces were selected based on the following considerations: percentage of children living with HIV, percentage of children on ART, number of COVida beneficiaries, number of COVida beneficiaries who were living with HIV, USAID priority status for a province, program stability, and security. Within each province, we selected 2 CBOs: 1 CBO with a high proportion of beneficiaries with unknown HIV status, and 1 CBO with a low proportion of beneficiaries with unknown HIV status. In each selected CBO, we randomly sampled 11 or 12 activists and then interviewed their managers (activista chefe) and supervisors.

### Data Collection Methods

Data were collected using qualitative and quantitative methods. From each CBO, 20 people were interviewed: 11 or 12 activists, 3 activist chefs (i.e., direct manager of activists), 2 supervisors (i.e., manager of activist chefe), and 3 other management staff. To be included, an activist must have been working for COVida for at least 6 months, and their activist chefe and supervisor had to be available for interviews. In total, 70 activists, 18 activist chefs, 12 supervisors, 6 CBO managers, 6 CBO monitoring and evaluation advisors, and the COVida project director were interviewed. Interviews were 30–45 minutes long, and all interviews were audio recorded and conducted privately at the CBOs. Data collection was done by a local research agency. Data collectors were selected based on level of education, prior qualitative interview experience, knowledge of the study areas, and fluency in study languages. A gender balance was ensured during recruitment of data collectors. Data collectors received 4 days of study-specific data collection protocol training. The data collection protocol was pilot tested and revised before conducting fieldwork. Data collection took place from July to August 2019.

Interviews elicited information including activista caseload, training, supervision, team meetings, non-monetary incentives, networking, demographics, work satisfaction, ways to improve service quality, time spent working, and any costs they incurred. Activista chefe and supervisor interviews also discussed challenges with activist retention, salaries, and activist performance. In interviews with CBO managers, advisors, and the project director, respondents discussed CBO-level procedures. Documentation was collected such as project reports, quarterly updates, and routine project data collected by COVida. Routine data tracked by COVida included activist caseload and beneficiary HIV status.

The study protocol received institutional review board approval from the Comité Nacional de Bioética para a Saúde in Mozambique and Health Media Labs, Inc. in Washington, DC. The informed consent process for interview participants was individualized and private. The data collectors privately shared information about the study with each potential participant and obtained and documented a written informed consent. Informed consent was administered in the language preferred by the participant. If consent was granted for audio recording, we recorded the interview. The information provided by respondents was held in strict confidence.

### Data Analysis

#### Preliminary Analyses

All interviews were translated and transcribed. The transcripts were qualitatively coded, a process whereby common and relevant themes are identified and sections of the transcripts that relate to the themes are tagged. The qualitative coding employed a deductive approach, in which topics related to modifiable case management attributes were identified by program experts and stakeholders before coding began. For example, “training” was identified as a potentially important
theme, because comprehensive training has been shown to positively influence program outcomes in health and resource-limited contexts. Microsoft Excel was used for the coding. Next, the coded data were reviewed, and summaries of each theme were created for each activista, activista chef, and supervisor. The summaries were compared across each of these 3 roles to identify and resolve conflicts. Conflicting statements were resolved by triangulating data from interviews and documentation, to ensure internal validity. The summaries were also used to identify differences between activistas across the modifiable attributes. This preliminary qualitative analysis was essential to set up the QCA.

Additionally, quantitative data from the interviews were analyzed. Descriptive statistics were calculated, using Microsoft Excel and SAS version 9.4, to summarize the range, mean, median, mode, frequency, and cumulative average of quantitative variables. These statistics were important to understand the spread of data across activistas for each modifiable attribute and to identify differences among activistas.

Qualitative Comparative Analysis
To identify the combinations of modifiable attributes that influenced knowledge of HIV status, a QCA was conducted. QCA combines quantitative and qualitative analyses to determine which combinations of variables (called pathways) influence the outcome analyzed. We selected QCA for this study since this method recognizes that several different combinations of variables may lead to a particular outcome. As a result, an implementing partner may choose any of the identified pathways to improve outcomes of interest. In QCA, the variables are referred to as causal conditions and are similar to independent variables in a statistical analysis. For this study, the modifiable attributes of case management are the causal conditions (e.g., caseload). The outcomes are similar to dependent variables and are the phenomena that are the main focus of the study. Because we did not want to lose information from data by restricting all conditions to dichotomous values, fuzzy-set QCA (fsQCA) was the analysis method selected, whereby fuzzy sets that ranged continuously from 0 to 1 were used to measure varying degrees of a condition’s presence or absence.

Outcome Identification and Calibration
Two outcomes were investigated for this study: (1) percentage change in knowledge of HIV status, and (2) percentage of beneficiaries with HIV status known at the time of the last assessment. Also, we conducted the analysis of the negation of the outcomes. We investigated the conditions that do not produce a high percentage change in knowledge of HIV status and conditions that produce high percentage of beneficiaries with HIV status unknown at the last assessment. The outcomes were calibrated using the direct calibration approach (see the QCA calibration guide in Supplement), in which a quantitative value associated with in-set membership, out-of-set membership, and the crossover point is first identified based on theory and the distribution of the raw, quantitative data; the data are then normalized between these points. Calibration is an iterative process between theory and collected data that aims to develop a common measuring stick to use to determine whether a case falls in the set of a phenomenon, out of the set of a phenomenon, or somewhere in between. The first outcome investigated was the percentage of an activista’s beneficiaries who changed their reported HIV status from unknown to known between enrollment and July 2019. These beneficiaries included all those who enrolled before April 1, 2019, and had their HIV status recorded at least 1 time in addition to the time of enrollment. HIV status was considered known if the beneficiary’s status was documented as HIV positive, on ART, not on ART (likely HIV positive but not receiving treatment), or test not recommended (likely not HIV positive, based on a risk assessment). HIV status was considered unknown if the beneficiary’s status was documented as unknown or not revealed. In-set membership was when the percentage of an activista’s beneficiaries with a change in HIV known status was greater than or equal to 75%, and out-of-set membership was when the percentage with a change in HIV known status was less than or equal to 25%.

The second outcome investigated was the percentage of an activista’s beneficiaries with their HIV status known at the last assessment. HIV status was considered known for beneficiaries whose status was either HIV positive or negative (and was not documented as unknown or not revealed) at the time of the last assessment. Beneficiaries for both second and third outcomes included all those who enrolled before April 1, 2019. In-set membership was when the percentage of an activista’s beneficiaries with HIV status known was greater than or equal to 95%, and out-of-set membership was when the percentage of an activista’s beneficiaries with HIV status known was less than or equal to 75%. 
Identification and Preliminary Removal of Potential Causal Conditions

A list of 23 potential causal conditions that were modifiable case management attributes was assembled from literature, case knowledge, and project documentation. Twenty-three causal conditions are too many for 70 cases in QCA because too much of the logic space or all of the possible combinations of conditions would not be represented by empirical cases. Conditions were removed based on lack of variation across the cases within a condition (known as domain conditions); correlations with other conditions, indicating that 2 conditions may be measuring the same topic; lack of data; or low necessity, indicating that the condition was less important for the outcome. Necessity is a QCA metric used to analyze individual conditions and can be helpful to narrow down a large list of potential conditions. Necessity reflects how important a condition is for an outcome, based on how often the condition is present when the outcome is present; necessity scores between 0.9 and 1.0 indicate that a condition is almost always necessary for the outcome.

Sufficiency reflects the extent to which a condition contributes to the presence of the outcome; a sufficiency score above 0.8 is required for a condition to be sufficient alone to produce the outcome. Necessity and sufficiency scores are produced by the software at the analysis stage.

We provide both necessity and sufficiency scores in Figures 1–3. As a check on the completeness of each final solution, the removed conditions were added back in and never resulted in higher solution consistency or coverage. From these initial analysis steps, the number of causal conditions analyzed was reduced to 11, which is a reasonable number for 70 cases using QCA: caseload, complexity, challenges in recruiting and retaining activistas, how cases are assigned, level of supportive supervision, out-of-pocket costs, quality of care team meetings, supervision ratio, time spent per case, training, and work experience.

Calibration of Causal Conditions

Two methods were used to calibrate the causal conditions. First, the indirect calibration method was used for conditions that primarily had qualitative data: challenges recruiting and retaining, how cases are assigned, level of supportive supervision, out-of-pocket costs, and training. Time spent per case was also calibrated indirectly, using primarily quantitative data. Project reports, documentation, literature, and expert knowledge were used to establish initial definitions for in-set membership (the criteria that correspond with a value of 1, when the condition is fully present for a given case or the case has full membership in the set of that condition), out-of-set membership (the

FIGURE 1. Outcome 1: Percentage Change in HIV Status Knowledge in Orphans or Vulnerable Children in Mozambique

Note: The numbering of the pathways is random; each pathway demonstrates an alternative combination of conditions that positively influenced the outcome and is considered to be equally sufficient in achieving the outcome.
FIGURE 2. Outcome 2: Percentage of Orphans or Vulnerable Children in Mozambique With HIV Status Known

FIGURE 3. Outcome 2 Negation Analysis: Percentage of Orphans or Vulnerable Children in Mozambique With HIV Status Unknown
criteria that correspond with a value of 0, when the condition is fully absent for a given case or the case has full nonmembership in the set of that condition), and the crossover point (the criteria that correspond with a value of 0.5, when the condition is neither present nor absent and is the point of maximum ambiguity). Next, qualitative data summaries were reviewed to determine meaningful differences between the activistas. For most of the indirectly calibrated conditions, 4-value fuzzy sets were used: 0 (fully out of the set), 0.33 (more out of the set than in), 0.67 (more in the set than out), and 1 (fully in the set); these sets are very common for fsQCA.32,34 For example, for the condition of challenges recruiting and retaining, in-set membership (a fuzzy set value of 1) was defined as:

The care team reports significant issues with recruiting and retaining activistas, is understaffed, and lacks a clear plan to recruit and retain activistas.

A value of 0.67 was defined as:

There are many issues with activista recruiting or retention, and activistas leave for reasons beyond the low subsidy. The care team may have plans to alleviate activista turnover, but no action has been taken.

A value of 0.33 was defined as:

There are some issues with activista recruiting or retention, such as activistas leaving due to low subsidies. The care team demonstrates clear actions and plans devised to alleviate activista turnover.

Out-of-set membership (a value of 0) was defined as:

The entire care team does not report issues with recruiting or retaining activistas, is fully staffed, and has a clear plan in place to recruit and retain activistas.

The remaining conditions, caseload and complexity, were calibrated using the direct calibration method, which is common for conditions with only quantitative data that can be normalized between anchor points.36 We measured complexity as the proportion of beneficiaries living with HIV in the activista’s case load. It was important to include complexity in our analysis since our interviews demonstrated that beneficiaries living with HIV often require more time from social workers to address issues such as comorbidities or treatment adherence, which would leave less time for other beneficiaries.

**Truth Table Assembly**

Once the causal conditions and outcomes were calibrated, the calibration criteria were used to assign fuzzy values to each case, for every condition and outcome. The qualitative coding and summaries and the quantitative values were used to determine whether each activista met the required criteria for in-set membership, out-of-set membership, or a membership value in between. These fuzzy values were assembled in a truth table that summarized the fuzzy scores assigned to causal conditions and outcomes for all activistas, reflecting the possible configurations of causal conditions associated with outcomes.14

**Truth Table Analysis**

Following the calibration of the causal conditions and outcomes, the truth table was analyzed using the “truth table analysis” function in fs/QCA software.38 Truth table analysis relies on the process of minimization, whereby stepwise comparisons between each combination of conditions are performed to determine which conditions could be removed to provide a more simplified pathway that consistently leads to the outcome.34,38 All possible combinations of conditions were investigated to determine (1) whether a given group of conditions consistently (i.e., nearly always) led to (i.e., was present when) the outcome occurred; and (2) whether the consistent combinations made sense with in-depth knowledge of the data (e.g., it would not make sense for challenges recruiting and retaining activistas to be a factor that contributed to a high percentage change in beneficiary HIV known status, but it would make sense for high-quality care team meetings to positively affect this outcome). Additionally, to further simplify the pathways, “easy” counterfactuals were used, called simplifying assumptions in QCA34; this action allows the researcher to specify assumptions for whether a causal condition’s presence or absence would be expected to be associated with each outcome. We made these assumptions based on case knowledge and theory of whether the presence or absence of a condition will lead to the outcome of interest. Thus, we assumed that presence of training will lead to high proportion of beneficiaries who know their HIV status. When there was not sufficient evidence to suggest the directionality of the condition’s influence on the outcome, both the presence and the absence of the condition were analyzed. For example, more time spent with each household could lead to a better outcome or it could lead to activista burnout and subsequently a poor outcome.
The validity of the results was determined based on 2 important QCA metrics: consistency and coverage. Consistency demonstrates the relative frequency that a pathway will result in a particular outcome or how consistently a pathway leads to that particular outcome; the accepted cutoff value for a consistent pathway is 0.8. Coverage is the percentage of cases with an outcome that is explained by a given pathway. Coverage lacks a cutoff value because it is a metric for generalizability; lower coverage scores, however, reflect more case-specific and less generalizable results. Once preliminary results were obtained, a subset/superset analysis was performed for each outcome to further simplify the number of causal conditions in each pathway while maintaining or increasing each pathway’s consistency and coverage. Finally, pathways were compared with theory and case knowledge to ensure that the final solutions presented the most complete and simplified explanations for the outcomes analyzed.

## RESULTS

### Study Participants

We interviewed 70 activistas, 18 activista chefes, 12 supervisors, and 6 CBO managers. The majority of activistas (71.4%) were women. The mean age was 30 years (standard deviation [SD]=9, range=20–57 years). Half of the activistas served beneficiaries in rural areas (51.4%). The majority of activistas (72.9%) had secondary education, one-quarter (24.3%) had primary education, and 3% had technical or professional education. Activistas’ total time working ranged from 6 months to 8 years, and 82.86% had 2 years of experience or less.

More than half of the activista chefes (55.6%, N=18) were women. The mean age was 34 years (SD=8.3, range=21–51 years). The majority of activista chefes (77.8%, N=18) had secondary education, less than a fifth (16.7%) had primary education only; 6% had technical or professional education. The number of years of work as activista chefes ranged from 6 months to 3 years. The majority of supervisors (91.7%, N=12) were men. The mean age was 40 years (SD=10.5, range=28–58 years). Two-thirds of supervisors (66.7%, N=12) had secondary education, and one-third (33.3%) had a university degree. The number of years of work as supervisors ranged from 6 months to 2 years. Lastly, 4 of the 6 CBO managers were women. The mean age was 51 years (SD=16.5, range=27–66 years). All 6 respondents had worked as CBO managers for 2 years.

### Outcome 1: Percentage Change in HIV Status Knowledge

For Outcome 1, 6 pathways led to a high percentage change in beneficiaries’ HIV status knowledge (Figure 1). Each pathway is 1 branch, with 5–8 conditions. For example, from Figure 1, Pathway 1 contained the following conditions: how cases are assigned, training, lack of challenges recruiting and retaining, complexity, quality of care team meetings, and less time per case. The solution consistency was 0.85 and coverage was 0.35. These pathways described the conditions that led to improved knowledge of HIV status for activistas from CBOs 2, 3, 4, and 6. Although some of the pathways are complicated (i.e., with more than 5 conditions), the solution reflects the most simplified combinations of conditions that consistently led to a high percentage change in HIV known status. No single condition was necessary or sufficient; instead, the combinations of conditions had the most important influence on the outcome.

#### Common Conditions

One condition was shared by all 6 pathways: how cases are assigned. This finding demonstrates that decisions to assign new cases should consider activista caseload, work experience, skills, case complexity, and activista proximity to case to prepare activistas well for effective case management. In particular, consideration of caseload and complexity of cases ensured that activistas were not overburdened and had enough time to address the needs of each beneficiary. In contrast, activistas who did not achieve high percentage changes in HIV known status were part of programs that lacked a formal procedure to assign cases. For example, an activista chefe from CBO 5 stated:

> The communities are divided into boroughs and the cases are allocated randomly.

Two other conditions were present in 5 of the 6 pathways: lack of challenges recruiting and retaining and lack of out-of-pocket expenses. For most activistas who achieved a high change in HIV status knowledge, their supervisors actively worked to retain activistas by ensuring that they had access to follow-up trainings and the opportunity to discuss job stress and complex beneficiaries. In CBOs 1 and 5 that had low percentage changes in HIV status knowledge, activista retention was inhibited by dissatisfaction with the activista stipend. In fact, two-thirds of activistas were unsatisfied with their stipend amount.
Activista Support and Experience
Five of the 6 pathways had at least 1 condition that reflected activista skills and external support: level of supportive supervision, quality of care team meetings, low supervision ratio, long training duration, and/or work experience (≥12 months). These pathways had a high amount of redundancy between conditions that ensure activistas were well equipped and well supported, which explains why most activistas were able to spend less time per case and still achieve desirable outcomes. Supportive supervision was the most effective when activista chefs met with each activista at least twice per week. Activista chefs traveled with all activistas periodically to:

know what is happening in the communities with the activistas, know the situations and look for joint solutions, [and] verify that the beneficiaries are receiving the services required from the activistas.

Low supervision ratios were not necessary for high levels of supportive supervision, but low ratios helped to ensure activista chefs and supervisors had more long-term time and energy to assist with complex cases.

Further, these results demonstrate alternative ways to support activistas, which may be useful to CBOs that have limited resources. For example, in Pathway 1, activista support occurred through training and high-quality and weekly care team meetings, where activistas shared experiences, discussed challenges, and created case plans. In Pathway 2, activista support occurred through training and because activistas had 2 or more years of case management experience. If a CBO has difficulty implementing regular care team meetings, hiring highly experienced activistas could provide an alternative. Similarly, training provided another layer of redundancy to activista support. Although the number of training days proved to be less important, the most effective trainings were those that employed multiple training activities (e.g., lectures, role playing, tests) and covered topics such as how to refer beneficiaries to health services, techniques to encourage ART retention, and how to cope with work stress. Overall, the presence of activista support conditions (e.g., quality of care team meetings, supervision ratio, supportive supervision) in multiple pathways may reflect how these conditions, in combination with other conditions (e.g., quality care team meetings and low supervision ratios) enable activistas to discuss difficult cases, which could help alleviate challenges of case complexity, for example), provided activistas with the resources needed to effectively manage their cases.

Complex Cases and Activista Preparation
The first 3 pathways show how activists who had a high percentage of complex cases that require more time still improved HIV status knowledge. The presence of high-quality team meetings, adequate training, and formal and thoughtful case assignment procedures prepared these activists well to manage complex cases. Finally, the presence of more time per case in Pathway 2 and less time per case in Pathway 1 may mean that this condition is not particularly important; activists can still achieve a high percentage change in HIV known status regardless of how much time they spend with each household. Some activists spent more time per case because they managed a higher proportion of complex cases; others spent less time per case because they had fewer complex cases or because their work experience or training meant that they strategized their cases to maximize efficiency. For example, 1 activista stated:

I first finalize my work at home and then plan my work, and I do this on a weekly basis.

Outcome 2: Percentage of HIV Status Known
The second outcome investigated was the percentage of beneficiaries with their HIV status known at the time of the last assessment. Five pathways led to a high percentage of beneficiaries with HIV status known (Figure 2), each with 4–6 conditions. This solution had a consistency of 0.87 and coverage of 0.42, which are acceptable values for QCA.34 This solution described activists from CBOs 2, 3, 4, and 6. No conditions were necessary or sufficient; instead, the combinations of conditions had the most important influence on the outcome.

Activistas With More Work Experience
Pathways 1, 2, and 3 shared 4 conditions: work experience ≥12 months, complex cases <10%, lack of out-of-pocket costs, and less time per case. The presence of work experience ≥12 months combined with complex cases <10% is one of the main reasons why activists were able to spend less time with each household (approximately 25 minutes or less, on average). In these first 3 pathways, activists worked efficiently and had lower risk of becoming burned out, allowing them to provide more effective case management services to beneficiaries. Additionally, these activists

Results demonstrated alternative ways to support activistas, which may be useful to CBOs that have limited resources.

Five pathways led to a high percentage of beneficiaries with HIV status known, each with 4–6 conditions.
had minimal challenges: lack of out-of-pocket costs and complex cases <10% meant that activistas had the resources to complete their work and did not have many cases that required more time. For activistas who did not have high percentages of beneficiaries with HIV status known, the presence of out-of-pocket costs could be a reason. These activistas spent their own money for transport, buying food for beneficiaries, buying airtime. If activistas have access to more resources, their effectiveness may improve.

For Pathway 1, the high supervision ratio meant that the ratios of activistas to activista chefe and of activista chefe to supervisor were higher (10:1 and 5:1, respectively). However, since these activistas had more experience and fewer case challenges, the supervision ratio did not negatively affect their case management effectiveness. In Pathway 2, activistas had caseloads of 50 or fewer cases and underwent significant training (>10 days); these conditions contributed to activistas’ preparedness and energy for effective case management. In Pathway 3, short training duration reflects fewer than 5 training days. Most training, regardless of duration, still covered important case management topics (e.g., goal setting, common challenges). Activistas requested additional training on topics including HIV testing, dealing with sensitive cases, and family planning.

**Activistas With Less Work Experience**

Pathways 4 and 5 differed from the first 3 pathways mainly because they had work experience ≤11 months. These pathways demonstrate 2 alternative combinations of conditions that are sufficient to overcome an activista’s lack of work experience. Notably, when an activista had less than 11 months of experience, it was essential that the activista had a caseload of 50 or fewer cases and that the activista attended weekly care team meetings that were comprehensive and addressed care issues beyond paperwork. For activistas managing more cases, especially those with less experience, external support and being well networked were important. Although activista chefe and supervisors identified the “ideal” caseload to be approximately 50 cases, this is effectively the maximum number of cases an activista should manage. For this dataset, activistas faced difficulty completing their case work when their caseload exceeded 50 cases, especially since they were expected to work only 20 hours weekly.

Although these activistas had less work experience, they received support that ensured they managed cases effectively. High-quality and weekly care team meetings with near-perfect activista attendance were essential to effective case management. In the highest-quality meetings, activistas were:

... Presenting the work [and] questions we have to the activista chefe who accompanies us to the beneficiaries’ houses [and] checking whether the forms have been filled well.

Support in developing weekly plans and sharing of information enabled activistas to gain important skills to manage difficult cases and was a valuable source of accountability for activista preparedness. In contrast, activistas who did not attend or did not have access to high-quality care team meetings described the meetings as unhelpful where the focus was solely for correcting forms, and this was associated with lower percentages of beneficiaries with HIV status known.

**Analysis of the Negation of the Outcomes**

The analysis of conditions that do not produce a high percentage change in knowledge of HIV status (i.e., negation of Outcome 1) did not produce any results. The analysis of negation of the Outcome 2 identified 1 pathway that led to a high percentage of beneficiaries with HIV status unknown (i.e., unknown or not revealed) (Figure 3). In-set membership was when the percentage of an activista’s beneficiaries with HIV status unknown was greater than or equal to 95%, and out-of-set membership was when the percentage of an activista’s beneficiaries with HIV status unknown was less than or equal to 75%. This pathway described activistas from CBOs 5 and 6 who had higher percentages of beneficiaries with unknown HIV status than the other CBOs. The pathway had 5 conditions: challenges recruiting and retaining, lack of ideal caseload, lack of supportive supervision, less time per case, and complex cases ≥10%. The solution consistency was 0.80 and the coverage was 0.09. No conditions were necessary or sufficient; instead, the combinations of conditions had the most important influence on the outcome.

**Activista Retention**

Challenges recruiting and retaining activistas had a negative influence on the percentage of beneficiaries with unknown HIV status. Activista chefe and supervisors attributed these challenges to low subsidies (i.e., activista salaries), delays in subsidy payment, lack of motivation, high caseloads, and

Global Health: Science and Practice 2020 | Volume 8 | Number 3
lack of job preparedness. For example, a supervisor stated:

Activistas leave because of the low subsidy or because of the workload in terms of caseloads.

Activista retention challenges led to higher activista turnover, which may have negatively affected case management.

**Activista Overwork**

The lack of ideal caseload means that activistas who had a high percentage of beneficiaries with unknown HIV status also had a caseload that was above the ideal/maximum caseload (n=50, based on the average number of cases that activista chefs and supervisors said an activista should ideally manage). The highest caseload cited was 106 clients. Too many cases meant that an activista had too much work and could not dedicate adequate time to each beneficiary, leading to an inability to know the HIV status of each. In contrast, activista chefs whose activistas had low percentages of beneficiaries with unknown HIV status aimed to assign activistas only the ideal caseload to:

...Not overload [them] with work [since] increasing the number [would] make it difficult for the activista to cover all the families.

In addition to too many cases, activistas with high percentages of beneficiaries with HIV status unknown additionally managed a higher percentage of complex cases, meaning they were often unable to spend as much time employing strategies to learn the HIV status of their remaining clients. Similarly, too many cases meant activistas spent less time per case, with some activistas spending as few as 15 minutes per household.

**Inadequate Activista Support**

Activistas with high percentages of beneficiaries with unknown HIV status also lacked supportive supervision. For example, regarding the purpose of the meetings with their activista chefe, an activista stated:

The activista chefe corrects the filled forms and we sign the central registry form.

In contrast, activistas who received highly supportive supervision discussed their difficult cases and created goals and plans during these meetings. For example, an activista stated:

I develop a plan, and if someone has abandoned ART treatment, that is where I need the activista chefe so that we can work together in this case.

Another activista summarized the tasks accomplished during these meetings as:

Talking about challenges, presenting the filled templates and being corrected and taught how to fill them correctly, undertaking simulations of the daily activities, weekly reporting, and exchange of experience.

**DISCUSSION**

This study aimed to provide actionable evidence to USAID and implementing organizations for strengthening their OVC programs. Previous studies have hypothesized ways to improve HIV case management, but there has been insufficient evaluation of the impact of these hypotheses on knowledge of HIV status, as well as their collective impact. The main findings from this study demonstrate the importance of strengthening multiple aspects of case management simultaneously, which may require efforts from both program funders and implementing organizations.

To achieve a high percentage change in knowledge of HIV status and a high percentage of beneficiaries with HIV status known, the implementing organizations working in a context similar to this study may select any of the identified pathways and work on ensuring the presence of each of the conditions in the pathway. The selection of the pathway may depend on the resources that are available and the timeframe that each organization has to work with. Thus, changing and implementing the procedures for assigning cases may be more time consuming than changing the caseload. Although we acknowledge that the interventions should focus on modifying all conditions in the selected pathway collectively, since the pathways as a whole were identified to influence the outcomes (as opposed to individual conditions), we would like to share several key recommendations regarding those conditions that were present in multiple pathways. First, it is important for programs to implement formal protocols that assign cases based especially on caseload, complexity, and activista skills. To our knowledge, how cases are assigned has rarely been mentioned in the literature, let alone tied to HIV program performance. Potentially, the biggest impact of formal protocols to assign cases was that they resulted in strategized case management (i.e., more experienced activistas managed the more complex cases) and reduced overwork. Caseload was important for the 2 positive HIV program performance outcomes investigated, so formal protocols could help to maintain ideal activista caseloads. For all activistas to have adequate time and
energy to manage their cases well, it is important that they do not manage more than 50 cases and manage a very low percentage of complex cases. Other programs have identified the maximum to be even lower (e.g., 30 cases), and programming is increasingly shifting focus to highly vulnerable populations that present more complex cases; therefore, the maximum should depend on how many hours an activista is expected to work weekly, their previous work experience, and how many complex cases they manage. This is particularly important for activists who have less prior work experience, since case manager expertise has been linked to improved HIV program performance.

Implementing organizations and program funders may also consider increasing material resources provided to activists because stipend increases may improve satisfaction and reduce activista turnover. Activistas with high percentage changes in knowledge of HIV status generally did not have out-of-pocket expenses. This finding is also supported by literature that suggests case workers are more effective when they have access to all resources essential to basic care. Adequate care resources and comprehensive care plans have also been shown to be important for treatment program effectiveness for people living with HIV. Program funders could consider expanding HIV program budgets to account for activista reimbursement for transportation and airtime (i.e., cell phone service used for case management). HIV treatment is also often inhibited by the lack of beneficiary access to food, so many activists mentioned that they purchased food for beneficiaries, resulting in out-of-pocket expenses. HIV program performance outcomes could further improve if program funders accounted for food for medication provision, which remains an unmet need.

HIV program performance may also improve if implementing organizations can provide activists with more than 1 form of high-quality support, especially to activists with more complex cases and/or less work experience. Support can occur through regular trainings, weekly care team meetings, and weekly individual meetings with activista chefs. In particular, high-quality care team meetings were important for activists who had less work experience since the meetings provided informal training and the opportunity to discuss challenging cases. Campbell et al. highlight the importance of social spaces for dialogue and critical thinking around HIV case management; regular team meetings can create this space and improve care team networking. Regular meetings with activista chefs could help activists to develop and be held accountable to comprehensive case management plans, which were often missing for activists with a high percentage of beneficiaries with HIV status unknown. These recommendations are supported by other studies that suggest that external support structures and the competency of management are critical to improve HIV outcomes.

Finally, programs could also implement protocols that seek to improve activista retention. Activista retention challenges were a major factor that inhibited knowledge of HIV status, often due to dissatisfaction with stipends and overwork. Dissatisfaction, low motivation, and emotional stress, have all been demonstrated to inhibit HIV outcomes. Activista retention challenges could be mitigated by increasing activista stipends and ensuring that they receive adequate training and supervision support.

Limitations

Work satisfaction, salary amount, care team networking, and strength of wider referral network were domain (constant) conditions, and therefore their influences on the outcomes could not be analyzed, presenting an opportunity for future research. Although all activists, activista chefs, and supervisors were highly or somewhat satisfied, programs could potentially improve satisfaction by increasing stipends and by using nonmonetary incentives to recognize good work. Nonmonetary incentives in other HIV treatment programs have helped to avoid demoralization and thus improve program performance. Respondents indicated that care team networking and the strength of wider referral network were strong. Literature highlights the importance of case worker connectedness to share challenges and of connections to additional services to ensure that beneficiaries receive comprehensive, whole-person care.

The influence of care team networking and referral networks is an important area for future research. Similarly, since CBOs all followed similar protocols for case management, the influence of specific care methods could not be investigated. For example, program performance may also be improved by increasing the mobility of care, such as using home-based rapid testing or performing follow-up consultations by phone. Future research that compares programs that employ different HIV testing methods and follow-ups and have different levels of care mobility could generate greater insight into program optimization.

Case management attributes are not the only influencers of knowledge of HIV status; however,
these factors were the focus of this study to identify actionable recommendations. Program performance may improve even more if the following individual-level factors can be addressed, which have been demonstrated to impact HIV outcomes: stigma,17,43 gender,22 urbanicity,22,44 beneficiary income,11,22 beneficiary education,22,23 and HIV testing method.22,42 We were also unable to consider different clinical and community-level factors that affect HIV outcomes, and these factors could potentially further improve HIV program performance outcomes. The following factors have been suggested in the literature to improve case management effectiveness and warrant further study: actual and perceived health service quality,17 distance to functioning health care facility,5,6,20,23 effectiveness of family support plans,11 new testing strategies (such as HIV self-testing, a focus on testing index cases of adults with HIV, and the availability of community-based testing), availability of HIV testing kits,17 and buy-in for HIV case management programs from the local community13 and district and central governments and nongovernmental organizations.6,15 Future research should seek to understand the intersection of case management attributes, beneficiary demographics, and clinical and community-level factors.

Lastly, this study is based on findings from 6 CBOs that were not selected randomly. Although we selected activistas within each CBO using simple random sampling, the results of the study may not be generalizable to other CBOs that are a part of the COVida project. Furthermore, this study relied primarily upon interviews with activistas and their managers to understand the important case management attributes. Although activistas estimated the amount of time they spent with each beneficiary and working for COVida per week, researcher observations could be important to better understand the relationship between time spent per case and HIV program performance outcomes. Similarly, study respondents described the importance of correctly and regularly completing case management paperwork, but paperwork completion itself was not a factor explicitly evaluated. Paperwork completion, follow-up phone calls, and other routine activist work could be included in a future study to identify additional ways to optimize HIV program efficiency and effectiveness.

CONCLUSION

QCA was used to investigate the combinations of modifiable attributes of HIV case management programs to understand how to improve knowledge of HIV status. Two outcomes were identified to measure case management effectiveness: percentage change in knowledge of HIV status, and percentage of beneficiaries with HIV status known at the time of the last assessment. We identified 6 pathways for the first and 5 pathways for the second outcome of interest. Each pathway demonstrates an alternative combination of conditions that positively influences the outcome. Implementing partners working in similar contexts as this study may select any of the pathways, based on their available resources, to improve the outcome since each pathway is equally sufficient in achieving the outcome. Overall, based on the presence of some of the factors in multiple pathways to improve knowledge of HIV status, we suggest that implementing organizations, donors, and governments focus on several key recommendations. Implementing organizations could implement a formal process to assign cases based on complexity and caseload to alleviate overwork. Implementing organizations could also aim to increase the level of support to activists, with activists having adequate external support such as high-quality care team meetings, direct managers who meet with activists weekly to assist with challenges, and/or low supervision ratios so that managers are available and not overworked. More program funding may be required to meet the need for increased activist support at current program target levels. We also suggest that, when possible, implementing organizations hire experienced activists and provide all activists with regular follow-up trainings so they have the tools to address challenging cases and complicated issues. This study builds theory on the important tenets of an effective case management program and has the potential to improve knowledge of HIV status for OVC and other vulnerable populations.

Acknowledgments: The authors thank the following individuals whose contributions made this study possible: Nathaniel Lohman, Celio Vilichane, and Dionisia Matos from USAID/Mozambique for their ongoing support and guidance throughout the study. We acknowledge COVida/FHI360 for their support in reviewing data collection tools and sharing program data for the study. We thank Hayley Bryant, COVida chief of party, for her work in conceptualizing the study. We greatly appreciate the efforts of Caximo Caixim, COVida monitoring and evaluation specialist, to provide us with program data for the analysis. We acknowledge the local data collection organization, Maraxis, for collecting high-quality data within a short period and for helping to code and summarize data from transcripts for the QCA. We also thank the MEASURE Evaluation knowledge management team, based at the University of North Carolina at Chapel Hill, USA, for editing, design, and production help. Importantly, we thank respondents from COVida and the 6 CBOs who participated in interviews. Without their input, our assessment would not have been possible.

Funding: This publication was produced with the support from the United States President’s Emergency Plan for AIDS Relief and the United States Agency for International Development (USAID) under the terms of MEASURE Evaluation cooperative agreement AID-OAA-L-14-00004. MEASURE Evaluation is implemented by the Carolina Population Center, University of North Carolina at Chapel Hill in partnership with ICF International; John Snow, Inc.; Management Sciences for Health; Palladium; and Tulane University. Views expressed are not necessarily those of USAID or the United States government. Partial funding support to complete the manuscript was also provided by Palladium.
REFERENCES


34. Schneider CQ, Wagemann C. Standards of good practice in qualitative comparative analysis (QCA) and fuzzy sets. Comp Sociol. 2010;9(3):397–418. CrossRef

Competing interests: None declared.
Drivers of HIV Status Knowledge in Orphans and Vulnerable Children Programs

Peer Reviewed

Received: October 17, 2019; Accepted: July 22, 2020


© Davis et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are properly cited. To view a copy of the license, visit http://creativecommons.org/licenses/by/4.0/. When linking to this article, please use the following permanent link: https://doi.org/10.9745/GHSP-D-20-00311

En português

Análise Qualitativa de Comparação dos Determinantes Chave do Estado de Conhecimento do HIV nas Crianças Órfãos e Vulneráveis em Moçambique

Constatações Principais

• Usando uma análise de comparação com uma “fuzzy set” identificamos combinações de factores modificáveis que os programas de HIV que apoiam órfãos, crianças vulneráveis e as suas famílias podem agir para aumentar a proporção de beneficiários que conhecem seu estado sorológico.

Implicações Principais

Para melhorar o estado do conhecimento do HIV, os programas devem considerar a implementação de um processo formal para:

• Atribuir os casos considerando a complexidade do caso e o número de casos para reduzir o excesso de trabalho.
• Fornecer apoio de supervisão aos activistas
• Contratar activistas experientes
• Formar activistas para resolver os casos desafiadores
• Reembolsar activistas por despesas relacionado com o trabalho

SINOPSE

Em Moçambique, mais de um milhão das crianças vivem com HIV ou são vulneráveis devido ao HIV. Respondendo a esta crise, o Plano de Emergência do Presidente dos Estados Unidos para o Alívio da SIDA financia programas que apoiam as crianças órfãs e vulneráveis afectadas pelo HIV e suas famílias. Esses programas retêm assistentes sociais, conhecidos como activistas, que fornecem e encaminham os beneficiários aos serviços para aumentar o conhecimento dos beneficiários sobre o estado de HIV para melhorar a retenção aos cuidados de saúde dos que vivem com HIV. Para melhorar eficácia do programa, as organizações de implementação precisam entender como os diferentes atributos de gestão de caso afectam os resultados dos beneficiários. Usamos uma Análise de Comparação Qualitativa com uma série difusa (QCA) baseada em 119 entrevistas para identificar as combinações de atributos de gestão de caso que levam para (1) maior conhecimento do estado de HIV e (2) altas percentagens de resultados dos beneficiários. Identificamos seis vias para o primeiro resultado e cinco vias para o segundo resultado. Cada caminho demonstra uma combinação alternativa de condições que influenciam positivamente o resultado e é igualmente suficiente para atingir o resultado.

Para melhorar o conhecimento do estado do HIV, os programas em um contexto semelhante ao deste estudo devem identificar as combinações de atributos de gestão de casos que levam para (1) maior conhecimento do estado de HIV e (2) altas percentagens de resultados dos beneficiários. Usamos uma Análise de Comparação Qualitativa com uma série difusa (QCA) baseada em 119 entrevistas para identificar as combinações de atributos de gestão de caso que levam para (1) maior conhecimento do estado de HIV e (2) altas percentagens de resultados dos beneficiários. Identificamos seis vias para o primeiro resultado e cinco vias para o segundo resultado. Cada caminho demonstra uma combinação alternativa de condições que influenciam positivamente o resultado e é igualmente suficiente para atingir o resultado.

Para melhorar o conhecimento do estado do HIV, os programas devem considerar a implementação de um processo formal para:

• Atribuir os casos considerando a complexidade do caso e o número de casos para reduzir o excesso de trabalho.
• Fornecer apoio de supervisão aos activistas
• Contratar activistas experientes
• Formar activistas para resolver os casos desafiadores
• Reembolsar activistas por despesas relacionado com o trabalho

SINOPSE

Em Moçambique, mais de um milhão das crianças vivem com HIV ou são vulneráveis devido ao HIV. Respondendo a esta crise, o Plano de Emergência do Presidente dos Estados Unidos para o Alívio da SIDA financia programas que apoiam as crianças órfãs e vulneráveis afectadas pelo HIV e suas famílias. Esses programas retêm assistentes sociais, conhecidos como activistas, que fornecem e encaminham os beneficiários aos serviços para aumentar o conhecimento dos beneficiários sobre o estado de HIV para melhorar a retenção aos cuidados de saúde dos que vivem com HIV. Para melhorar eficácia do programa, as organizações de implementação precisam entender como os diferentes atributos de gestão de caso afectam os resultados dos beneficiários. Usamos uma Análise de Comparação Qualitativa com uma série difusa (QCA) baseada em 119 entrevistas para identificar as combinações de atributos de gestão de caso que levam para (1) maior conhecimento do estado de HIV e (2) altas percentagens de resultados dos beneficiários. Identificamos seis vias para o primeiro resultado e cinco vias para o segundo resultado. Cada caminho demonstra uma combinação alternativa de condições que influenciam positivamente o resultado e é igualmente suficiente para atingir o resultado.

Para melhorar o conhecimento do estado do HIV, os programas devem considerar a implementação de um processo formal para:

• Atribuir os casos considerando a complexidade do caso e o número de casos para reduzir o excesso de trabalho.
• Fornecer apoio de supervisão aos activistas
• Contratar activistas experientes
• Formar activistas para resolver os casos desafiadores
• Reembolsar activistas por despesas relacionado com o trabalho
Matching Development of Point-of-Care Diagnostic Tests to the Local Context: A Case Study of Visceral Leishmaniasis in Kenya and Uganda

Michel Bengtson, a* Mitasha Bharadwaj, a* Astrid ten Bosch, b Hellen Nyakundi, c Damaris Matoke-Muhia, d Cees Dekker, a Jan-Carel Diehl b

Key Findings
- A concept target product profile is an effective new tool that can aid researchers to develop a technology to match specific health care context more quickly than a conventional target product profile.
- The role of local volunteers and community health care workers is critically important for access to diagnostics in resource-limited settings.
- A noninvasive test-of-cure and a screening and confirmation test will significantly improve management of visceral leishmaniasis in endemic regions.

Key Implications
- Early during the research and development stage, researchers should consider who will administer the test (patient, health care worker, doctor) and for what purpose.
- Program managers should consider that the training level of staff and volunteers and the availability of the resources are the critical determinants for the use of a diagnostic test.

ABSTRACT

The rapid growth of point-of-care (POC) diagnostic tests necessitates a clear vision of when, where, and why a new POC diagnostic test needs to be developed and how it can be used in a way that matches a local health care context. Here, we present an innovative approach toward developing a concept target product profile (CTPP), which is a new mapping tool that helps researchers match a new diagnostic test to a specific local health care context early in the research and development process. As a case study, we focus on the diagnosis of visceral leishmaniasis (VL) in rural resource-limited regions of Kenya and Uganda. Our stepwise approach integrates elements of design thinking and uses a combination of literature reviews and field research for a context analysis of local health care systems and practices. We then use visual thinking in the form of Gigamaps and patient journeys to identify use case scenarios and to present our findings from the field research to key stakeholders. The use case scenarios describe the diagnostic scope of a new POC test based on the feasibility of the new test, the local need, and the contextual fit. For our case study of VL, we identify 2 valuable use case scenarios, namely test-of-cure and screening and confirmation, and we formulate a CTPP. We anticipate that a CTPP will enable researchers to match a new POC diagnostic test during the research and development process to the local health care context in which it will be used.

INTRODUCTION

Within the past decade, point-of-care (POC) diagnostic tests have received immense attention because their accuracy and ease of use create the ideal solution for early diagnostics of infectious diseases in resource-limited settings. POC diagnostic tests should involve a minimum number of steps to obtain a real-time result that is easy to interpret. Importantly, POC tests are designed to be performed near or at the site where the patient is to enable a short turnaround time for them to receive treatment and care. Given these favorable characteristics, novel POC technologies for decentralized diagnostics are being developed worldwide at a rapid rate. Ample research opportunities exist for continued development of POC diagnostic tests, particularly for neglected tropical diseases (NTDs), which are a group of...
chronic, disabling, and potentially fatal diseases that are prevalent in tropical and subtropical regions. NTDs occur predominantly in resource-limited settings, which are defined by the World Bank as resource-constrained (human, environmental, economical) regions with limited infrastructure and/or basic services in low- or middle-income countries. By definition, NTDs receive little attention, and 14 of the 24 NTDs that are currently acknowledged by the World Health Organization (WHO) lack essential POC diagnostic tests.

It is crucial to ensure that new POC diagnostic tests meet the needs of the (multiple) end users and fit constraints within local health care contexts. Particularly for resource-limited settings, innovative approaches are required to meet the demands for diagnostics that are suitable for use at the lowest and most constrained level of the health care system. Current POC diagnostic tests are often not compatible with the most constrained level of the health care systems in resource-limited settings because they still require resources, such as a cold chain to store reagents or trained users who are generally not available in such settings. Although these resources are available at higher levels of the health care system (reference laboratories and hospitals), they are not available in the local clinics in remote areas where patients first seek health care. This situation necessitates new POC diagnostic tests that do not depend on additional resources and can be administered by users without extensive training.

Multiple guidelines are available that can be used to jointly guide the design of new POC diagnostic tests that address the needs of the end users and stakeholders in a local health care context. WHO developed the ASSURED criteria (i.e., a diagnostic test should be affordable, sensitive, specific, user-friendly, rapid and robust, equipment-free, and deliverable to end users) to guide the development of new medical technologies and to encourage the adaptation of existing technologies to suit resource-limited settings better. However, the ASSURED criteria are rather general and broadly applicable. WHO and other organizations, such as the Drugs for Neglected Diseases initiative and the Foundation for Innovative New Diagnostics, collaborate with various stakeholders and experts to develop target product profiles (TPPs), which are more specific guidelines that provide details on the minimal and optimal performance and operational features of diagnostic tests. TPPs are the end result of several rounds of discussions to reach a consensus from policy makers, which is an important yet labor-intensive process. TPPs are developed when the use cases are already defined and it is known when, where, and why the test will be used, that is, when specifications have been determined regarding the diagnostic moment (when the patient gets tested), the diagnostic setting (where the patient gets tested), and the purpose of the test (why the patient gets tested), which could be screening and confirmation.

While developing a novel POC diagnostic test, researchers would benefit from having guidelines that are less abstract than the ASSURED criteria, which may be too broad to be effective for a specific context and disease but are less involved and less prescriptive than a full TPP. Indeed, a POC diagnostic test that does not meet all the stringent requirements of a TPP can still be of great value, depending on the local health care context and the needs of the end users. For example, if a TPP defines a desired sensitivity of a test, but a new test is somewhat less sensitive (and thus would be excluded by the TPP) yet is much more stable at higher temperatures or an order of magnitude cheaper, it could still be valuable for a specific disease in an endemic tropical region. To facilitate the design of a POC diagnostic test that fits a health care context, we conclude that an efficient and effective mapping tool is needed, particularly when the diagnostic moment, setting, and purpose of the test are not yet defined.

To address this need, we propose to formulate a concept target product profile (CTPP) as an intermediate guideline for developing a diagnostic test that addresses the needs of the end users in the local health care context. A CTPP does not replace a TPP, and it best serves as an intermediate guideline for researchers in the form of a mapping tool. Such a CTPP should preferably be developed at the onset of the development of a new POC diagnostic test, aiming researchers in considering the context at an early stage while the POC diagnostic test is designed and developed. A CTPP would enable researchers to identify the minimum features for successful implementation of a POC diagnostic test in a local health care context, and it will cost considerably less time and resources than the development of a TPP. To illustrate, a recent TPP for dermal leishmaniasis took approximately 4 years and 82 experts to reach a consensus, whereas a CTPP as proposed in this study can take 6 months and a considerably smaller team depending on the interdisciplinary team of choice. Although a CTPP will serve a clear
purpose, a full TPP is still required because it is extremely valuable for experts such as policymakers and provides guidelines for further test development by technical experts. In this paper, we propose a step-by-step approach, which includes elements of design thinking, toward developing a CTPP that matches the new POC diagnostic test and the local health care context in which it will be used (Figure 1).

**METHODOLOGY**

For the development of a CTPP, we adopted a stepwise approach, using design-thinking principles (Box). Design thinking is a human-centered approach to innovation that integrates the needs of people (desirability), the possibilities of the technology (feasibility), and the requirements for business success (viability). Design thinking encourages novel and thorough solutions. In general, design thinking is utilized to evaluate the current situation, to identify the actual problem, and thereby provide a guideline for developing customized solutions. Core principles of design thinking are to empathize, define, ideate, prototype, and test. The first 3 of these principles have been integrated into our stepwise approach for formulating a CTPP.

In our case, the design-thinking principles were used to evaluate current diagnostic practices for a disease, identify the diagnostic need for a disease, and provide a diagnostic solution. The approach started with a comprehensive literature survey, followed by observations and semistructured interviews with the health care providers in the field and scientific researchers. In this step, the design principle empathize was integrated to gain an empathetic understanding of the problem at hand. Thereafter, the design-thinking principle define was utilized to combine the information gathered and identify the problem at hand. This step used a set of design tools known as “visual thinking.” Specifically, “Gigamaps” were used to describe the health care system and define the current diagnostic need. Gigamaps are large and information-dense diagrams that act as a bridge between inquiry, design, and implementation. Using these, researchers are intentionally encouraged to identify and subsequently use patterns that emerge from the field observations and data. Thereafter, visual depictions of the patient journeys were constructed based on the information gathered from literature and field research. The term patient journey refers to the experiences and processes that a patient goes through during the course of a disease and its treatment. These patient journeys provide a detailed yet simplified overview of the challenges faced by the patients while seeking effective diagnoses and subsequent treatment. Finally, the design-thinking principle ideate was applied to obtain the logical solution for the diagnostic needs defined in the previous step. Again, a visual thinking tool in the form of use cases, referred to as scenarios, was utilized to present the
A CTPP is expected to be broadly applicable for diagnostics in endemic resource-limited settings.

**BOX. Design-Thinking Principles**

Design thinking refers to a strategic process in which design concepts are developed (e.g., proposal for a new diagnostic test) before final product development. It is essentially an iterative method for problem solving. The use of design thinking encompasses the innovation of new products within social and business contexts. In this way, designers actively seek to understand the user and their experience before they try to solve the problem and move into the product development process. Therefore, design thinking involves methods such as context analysis, problem framing, creating thinking, prototyping, testing, and (re)evaluating.

Core principles of design thinking are to empathize, define, ideate, prototype, and test. Design thinking aims to create a solution that is desirable (for the end user and stakeholders), (technically) feasible, and (economically) viable. Within the broad context of design thinking, tools such as visual thinking can be used to facilitate the design-thinking process. Visual thinking allows people from diverse backgrounds to share insights, synthesize existing information, and formulate superior ideas. Visual thinking tools enable designers to convey complex ideas to the stakeholders in a logical manner. To summarize, design thinking is an unconventional toolbox that within its visual thinking framework can facilitate the communication of complex problems and thereby aid in inventing/devising novel solutions to complex problems.

complex problem and logical solution. Based on the outcomes from our approach, a CTPP was formulated that integrated a desirable, feasible, and viable solution to a complex societal problem.

The above methodology is expected to be broadly applicable in the field of diagnostics for use in endemic resource-limited settings, since context analysis is always essential to guide the design process during R&D. Developing a CTPP requires the following seven steps.

**Step 1: Literature Review**

A critical review of the existing literature on a disease and its relevant health care context is performed to gauge the disease endemicity, resource availability, and current diagnostic practices in the resource-constrained settings. Such critical analyses help in identifying the potential and the limitations of the existing diagnostic practices and in identifying stakeholders such as patients, health care staff, patient families, health care organizations, nongovernmental organizations, and government bodies, and thereby highlight the implementation needs for novel diagnostic solutions.

**Step 2: Selection of an Endemic Resource-Limited Region for a Case Study**

Beyond mere literature study, it is important to direct field research for a case study. Selection criteria for a fitting case should include (i) an endemic region that bears a high burden of a disease; (ii) an area that has an urgent need for POC diagnostic tests; and (iii) a politically stable conflict-free zone for safety and logistic capacity (roads and access to remote endemic areas) for field research.

**Step 3: Field Research With Direct Observations and Interviews With Stakeholders**

Despite an extensive literature survey, many aspects of diagnoses in remote endemic regions often remain unclear, such as the clinical algorithm for disease identification, the patient’s journey from becoming infected to getting treatment, the availability of laboratory equipment, and the level of trained personnel at the lowest most constrained level of the health care system. To obtain observations from the field for filling the knowledge gaps in the literature, a field trip to the selected endemic regions was necessary. The objective of such a field trip is to gather direct observations at various health care levels in that region and to carry out semistructured interviews in the field to obtain expert input from key stakeholders as identified from the literature survey and from advice of the locals.

**Step 4: Create Gigamaps and Patient Journeys Based on Insights Gathered**

To get a deeper understanding of the diagnostic practices for a disease, the assumptions made by the detailed literature survey (step 1) are analyzed in conjunction with the information that was obtained from semistructured interviews with the health care professionals (step 3). The observations are processed and organized to highlight the limitations of existing diagnostic practices. Since the health care context of a disease in resource-limited settings is complex with social and technical challenges and involves a wide range of stakeholders, understanding the corresponding diagnostic practices for a disease can be quite challenging. To deal with the complexity of such a multifaceted health care system, visual thinking tools are used to present the complex diagnostic problem at hand (define) and to identify the logical solution (ideate) in the form of use case scenarios. Visual Gigamaps are created to outline current diagnostic practices at various stages of the disease, and health care-seeking behaviors of patients from a health care
provider’s point of view across different health care levels. Next, visual patient journeys are constructed within the selected disease endemic region. In our approach, these extended patient journeys represent a sequence of interactions between the patient, the health care system, and the stakeholders involved, and they consider both the technical, economic, and social factors.

**Step 5: Create Use Case Scenarios**
Next, we create various scenarios for the application of a diagnostic test within the selected region. Again, visual thinking is utilized to present every scenario to provide a clear understanding of the requisite diagnostic need and to suggest logical solutions.

**Step 6: Validation of the Use Case Scenarios**
Visualizations of the scenarios from step 5 are used to facilitate discussions with different stakeholders with an objective to critically select and thus validate the most valuable scenario through which a diagnostic test can meet the local needs of the end users and stakeholders. Thus, detailed discussions are conducted with the stakeholders on each scenario to obtain the most pressing diagnostic need and to define the priorities from the perspective of health care providers.

**Step 7: Formulation of a CTPP**
Finally, the stepwise approach collectively leads to the formulation of a CTPP for a diagnostic test that will be most effective for the selected disease-endemic region. The CTPP is formulated for the most urgent diagnostic need as viewed from the perspective of health care providers.

### RESULTS OF A CASE STUDY: DEVELOPING A CTPP FOR A POC DIAGNOSTIC TEST FOR VL IN KENYA AND UGANDA

For proof of principle, we selected the NTD visceral leishmaniasis (VL) as a case study to validate our CTPP design approach. Below we explicate the above 7 steps for this case study.

#### 1. Literature Review on VL Endemicity and VL Health Care Context
Leishmaniases are caused by more than 20 different *Leishmania* species, which are parasites that can be transmitted to humans and other animals by the bite of infected female phlebotomine sand flies. Worldwide, there are approximately 2 million new cases each year, and 556 million people are at risk of acquiring the infection. VL (also known as Kala-azar) is the second-largest parasitic killer after malaria and the most severe form of leishmaniasis because it affects the visceral organs, particularly the liver, spleen, and lymph nodes. Although VL is curable, it remains a fatal disease because it is often left untreated due to its low index of suspicion by health care providers, late diagnosis, and inadequate case management, especially at an early stage in low-resource settings. Furthermore, the initial symptoms of VL (e.g., persistent fever, weight loss, fatigue, and anemia) overlap with other febrile illnesses, such as malaria, and hence it is often misdiagnosed and treated incorrectly. If left untreated, VL is fatal within 2 years, due to severe anemia or secondary bacterial infections. Furthermore, post-kala-azar dermal leishmaniasis (PKDL) is a skin condition that occurs after VL treatment due to persisting parasites in the skin. Incomplete treatment is a major risk factor for PKDL. Although PKDL lesions are typically self-healing, they pose as infectious reservoirs for sandflies and cause aesthetic and psychological complications that affect the patient’s quality of life, especially young adults. VL is a complex disease that is further complicated by coinfections such as HIV. Individuals with HIV who are immunosuppressed often present with more severe VL symptoms and require different treatment regimens. Other common coinfections in VL endemic regions include malaria and tuberculosis. Thus VL needs to be diagnosed and managed on a case-by-case basis due to confounding conditions such as immunosuppression.

VL is currently diagnosed by using either one or a combination of the following: (1) empirical clinical observations, (2) immunological rapid diagnostic tests (RDTs) and/or immunoassays, (3) molecular tests to detect the pathogen’s DNA in clinical samples, and (4) parasitological tests that require microscopic analyses of invasive splenic or bone marrow aspirations. Figure 2. Except for clinical methods and RDTs, which can be less reliable, these diagnostic practices require expensive instruments, a stable source of electricity, a well-equipped laboratory, and an expert to operate; therefore, they are inadequate for use within resource-limited settings. The most readily used test for VL diagnosis, especially in remote settings, is the rK39 RDT. However, since the rK39 RDT is an immunological test, it is less reliable because the sensitivity of the test differs between individuals.
and it cannot serve as a test-of-cure due to persisting antibodies after treatment.30,37

2. Select a Region for a Case Study
Leishmaniases remain endemic in more than 98 countries with the majority of VL cases in South Asia (India, Nepal and Bangladesh), South America (Brazil), and the horn of Africa (Ethiopia, Somalia, South Sudan, Sudan, Kenya, and Uganda) (Figure 3).31,38 Seven countries in particular (India, Brazil, Ethiopia, Kenya, Somalia, South Sudan, and Sudan), reported approximately 90% of the global cases of VL in 2015.29 Although various VL control programs that focus on prevention and treatment are operational within these countries,22 the current rK39 RDT has been shown to have a poor performance in East Africa compared with India,29,36 necessitating an improved diagnostic approach in East African countries.

Considering the global spread of VL, our first challenge was to select a VL endemic region that was conducive for field research. Based on our
selection criteria (see Methodology), we selected western Kenya and northeastern Uganda for our field research.

3. Field Research With Direct Observations and Interviews With Stakeholders

During our field trip to western Kenya and northeastern Uganda, we visited several Pokot tribal communities, health care facilities, and local organizations over 2 weeks in November 2018 (Figure 4). Our international team consisted of scientific researchers including a principal investigator, a postdoctoral researcher, and a PhD researcher, who are working together to develop innovative POC diagnostic tests for infectious diseases, and an industrial-design master student. Our local Kenyan team consisted of a public health officer, a research technician, and a research assistant. We engaged with county and subcounty officials, as well as with local health administrators and community health volunteers (CHVs). Our Ugandan team consisted of a medical doctor and a community health worker (CHW), and we engaged with key stakeholders such as the local chief in the Moroto district. A significant number of interviews were conducted as we followed the recommendations and advice of the locals about whom to speak to (Figure 4).

In Kenya, VL testing and treatment can be accessed at the local health care facilities that are located in the Rift Valley region: the Kimalel health center and a newly constructed treatment facility at the Chemolingot subcounty hospital (both within the Baringo County), and the Kacheliba health center (within the West Pokot County). In Uganda, the Amudat hospital is the only VL treatment facility. VL testing can be accessed in Rupa subcounty and by a mobile CHW in the surrounding regions, and patients are referred to the Amudat hospital for further testing and treatment.

4. Create Gigamaps and Patient Journeys Based on Insights Gathered

A Gigamap of the health care system in western Kenya and northeastern Uganda was created based on the literature as well as on the insights obtained in the field. Seven phases were identified (Figure 5), from exposure to vectors at home (phase 1); being sick (passive) (phase 2); seeking care (active) (phase 3); getting diagnosed with VL (phase 4); getting to a treatment facility (phase 5); getting treatment (phase 6); and finally, to being treated and going home (phase 7). The Gigamap visualized the journey through the different levels of the health care system, starting from the rural setting (close to homesteads) and advancing towards more urban settings for treatment. The Gigamap also visualized the multiple stakeholders that are involved in the VL health care system.
Next, we created patient journeys (Figure 6). We related these journeys to the 7 phases defined in the Gigamap, to obtain a detailed overview of the challenges faced by a patient while seeking effective diagnoses and subsequent treatment. The 8 patient journeys (numbered as I to VIII) were a result of the information gathered in the field, and they represent a sequence of interactions between the patient, the health care system, and the stakeholders involved. A detailed description of patient story II can be found in the Supplement. Possible factors and barriers encountered during the patient journeys were identified. Several factors determined the progress of a patient through these different phases resulting in the least efficient (story II) and the most efficient (story VIII) journey from infection to the treatment.

An important observation during our study was the key role of the CHV/CHW as the closest link to the people in rural communities. In Kenya, CHVs and CHWs are trained on a variety of health issues including case definition, prevention, and control of common ailments, as well as nutrition and family planning. They facilitate access to health services through advocacy, outreach, referral, community education, informal mentoring, and social support. Thus, their role is to identify patients at the homesteads (small clusters of homes) (Figure 6, phases 1 and 2) and refer them to the local health care facilities for further diagnoses (Figure 6, phase 4). In Uganda, we observed that CHWs were trained to perform VL diagnostics in the field using the rK39 RDT. Due to ongoing VL clinical trials, which raises funding availability, CHWs in Uganda also had access to a motorcycle to allow them to reach secluded homesteads.

In contrast, due to limited clinical trials and a subsequent lack of funding, we observed that CHVs in Kenya are involved in identification of patients and referring them to local health care facilities. CHVs are well-respected members of the communities, such as teachers and ministers, and are trusted by the locals. Furthermore, due to language barriers between the locals and researchers from abroad, CHWs and CHVs played a key role in most interactions with the locals, patients, and health professionals.

The distance to a health care facility for diagnoses and treatment is the most significant barrier that patients face when seeking health care.
in Kenya, that may have VL based on the rK39 RDT, currently travel approximately 80 km to the Kimalel hospital for another (confirmation) diagnostic test and for treatment, if needed. Due to the toxicity and costs of the treatment, patients are diagnosed multiple times to ensure that treatment is only prescribed when absolutely necessary (in contrast to antimalarial treatment that is prescribed more readily). The Chemolingot sub-county hospital treatment facility in Kenya is currently improving accessibility to treatment for patients in the East Pokot subcounty. However, in Uganda, patients from across the country need to travel to Amudat for VL treatment. Interestingly, many Turkanas in Kenya travel approximately 100 km to the Amudat hospital to seek treatment due to conflicts between their tribe and the Pokot in Kenya.

Overlapping symptoms with other febrile illnesses often lead to misdiagnoses based on empirical clinical observations (signs and symptoms) (Figure 6, phases 2–7). In VL endemic regions, acute fever is often associated with malaria and other prevalent tropical diseases. Malaria RDTs and treatment are generally readily available, which encourages their use. Furthermore, a lack of VL awareness, even at the health care centers, promotes such misdiagnoses. Health education influences the ability of patients and health care workers to recognize VL. Increased VL awareness will have a positive influence on the journey of the VL patient because it increases the likelihood of recognizing VL at an early stage. Furthermore, we learned that children play in the termite mounds that are preferred resting and breeding sites for sandflies (the VL vector), and males sleep outside at night, which further exposes them to sandfly bites. Important risk factors include area ecology (humidity, heat), vegetation (acacia trees), livelihoods (pastoralism and proximity to livestock), and general behavior (outdoor sleeping).

A significant lack of resources is apparent for the diagnosis and treatment of VL, particularly in terms of health care personnel, diagnostic testing, and financial resources.
and financial resources. Despite traveling far distances, patients have no guarantee that diagnostic procedures will be practiced immediately once they arrive at a health care facility for diagnoses. We learned that technicians generally close their laboratories for a day or two when they need to be in the field or are receiving training. Due to a general shortage of staff, diagnostic procedures are often delayed. Furthermore, RDTs such as the rK39 for VL are not always available, which promotes misdiagnosis of VL. Midwives at the Rupa health center in the Moroto district in Uganda, who routinely perform RDTs for malaria and other illnesses, informed us that the supply of RDTs for VL is not consistent. Thus, symptomatic patients that do not respond to malaria medication are then referred to the Amudat hospital (120 km) for diagnoses and treatment. Finally, the cost of traveling to a health care facility is a significant barrier, particularly when patients are asymptomatic and/or are simply not convinced that they are sick due to a lack of awareness. In general, a lack of financial resources contributes to poor health-seeking behavior (Figure 6, phase 2).

We learned that many VL patients know that they are sick (Figure 6, phase 2), but do not actively seek health care (testing and treatment) due to the inaccessibility of health care facilities and the fear of encountering other tribes that may be hostile. We also learned that the Pokot and Turkana tribes in Kenya and Uganda are often in conflict, which makes it unsafe for members of either tribe to travel freely to seek health care. Such tribal conflicts inhibit patients from seeking health care and force them to travel to more distant health care facilities that are located away from a conflict zone. Such delays in seeking health care could further worsen the patient’s health, and they are often severely anemic and weak by the time they reach a health care facility. Furthermore, patients require screening for multiple infections, including VL, malaria, and tuberculosis, which is challenging when the patients are extremely weak. Stabilizing patients before starting them on VL treatment is also challenging because they may require blood transfusions and treatment for comorbidities, such as malaria, which require urgent attention before they can start VL treatment.
We observed that in general, VL patients also seek care from a traditional healer before considering visiting a health care facility because traditional healers are trusted, located closer to the homesteads, and often alleviate some of the initial symptoms. These deferrals worsen the patient’s health, thereby making the treatment more difficult. Unfortunately, critically ill patients often do not respond to the VL treatment, resulting in death. Such cases further strengthen the traditional beliefs and set a negative impression of health care facilities in the minds of health care seekers. Due to financial constraints and the cultural beliefs of the local population, the health care facility usually needs to arrange the last rites because the family does not come to the hospital to take responsibility for the deceased.

We observed a traditional patriarchal society, in which males leave the homesteads for work, and females take care of the homes and the children. For males, the loss of income due to the time that is spent traveling to seek health care contributes to poor health-seeking behavior (Figure 6, phase 2). For females, being unable to leave children unattended at home contributes to poor health-seeking behavior. We learned that women from the Pokot tribes often need permission from their husbands to leave their homes before taking themselves or a child to a health care facility. Thus, unequal decision-making power in the household contributes further toward delayed diagnoses and VL treatment for women and children.

Given all these factors, the total number of steps a patient takes to seek treatment can vary significantly. For example, seeking care in phase 3 requires many steps in complicated patient journeys, such as for patient stories I and II (Figure 6). This may be due to several of the aforementioned factors. For example, a lack of resources in the unavailability of staff or RDTs may result in the patient returning home (Figure 6, phase 2) without receiving a diagnosis, and later traveling back to seek care (Figure 6, phase 3). Traveling between phases 1 and 2 may be further hindered by other factors such as distance between the homesteads and the health care facilities or a lack of financial resources. Conversely, a patient journey may be as simple as patient story III whereby a patient progresses easily from being sick (Figure 6, phase 2) to being treated and going home (Figure 6, phase 7).

5. Create Use Case Scenarios
Prompted by our extensive methodology, whereby we mapped the health care system in a specific endemic region and analyzed patient journeys, we sketched 6 use case scenarios based on the scope of application of a new POC diagnostic test (Figure 7). These different scenarios are based on considerations involving 3 elements: the characteristics of the new diagnostic technology, the contextual fit, and the local need. Each of the scenarios describes and visualizes a potential specific health care context in which a test based on a new diagnostic technology could be of added value to improve diagnostics in the health care system. Thus, each scenario represents a potential diagnostic setting.

6. Validation of the Use Case Scenarios
The aforementioned 6 scenarios were discussed with experts in the field, including a Kenyan public health officer and VL specialists at Médecins Sans Frontières, to identify the most urgent need based on the perspective of health care providers. Thereafter, we chose a researcher-centric approach to obtain a context-specific diagnostic need that would facilitate researchers to outline the technological requirements of a new POC test. Interestingly, 2 scenarios—screening and confirmation and a test-of-cure—were consistently identified as a priority for developing a CTPP, and we did not obtain any discrepancies in the opinions of the experts. The selection of these 2 scenarios was based on the following criteria: how well the scenario represents a diagnostic setting in the current health care context of VL; how well the scenario meets a local need in terms of VL case management; and how feasible it would be to implement a new POC test in the scenario.

The first scenario selected, “screening and confirmation,” is beneficial for a number of reasons. Given the poor performance of the serological rK39 RDT in east African countries, a more specific and more sensitive POC diagnostic test that can be implemented by end users with minimal training at the lowest level of the health care system is clearly required. Furthermore, it is crucial to consider the end user(s) of a POC diagnostic test because the level of training and availability of resources will influence the diagnostic setting. Thus, a simple, noninvasive, yet effective VL POC diagnostic test that can be used for initial screening and confirmation (scenario 1) would significantly improve VL case management because it would replace the rK39 RDT for initial screening and invasive splenic aspirations for confirmation. The added value of an effective screening and confirmation test is that patients could be screened more reliably at the lowest level of the health care system, for example, by a CHV/CHW who is closer to the homesteads, which would prevent patients...
from traveling unnecessarily to regional health care facilities.

The second scenario, “test-of-cure,” is useful because it would replace cumbersome procedures (i.e., microscopic analysis of invasively obtained splenic aspirations or molecular tests such as polymerase chain reaction) that are currently used for test-of-cure. Serological tests cannot serve as test-of-cure owing to persisting antibodies after VL treatment. Thus, a POC diagnostic test that can serve as a test-of-cure (scenario 2) would significantly improve VL case management by replacing invasively obtained splenic aspirations. The added value of a test-of-cure is that relapse of the disease, which occurs in approximately 10% of VL patients, could be detected at an early stage after treatment. Therefore, the test-of-cure scenario fills a critical gap in VL case management.

Scenarios 3 to 6 were excluded from further development of a CTPP. A screening day, as depicted in scenario 3, would be challenging to implement because one cannot ethically test everyone in a community, especially not asymptomatic patients, while symptomatic patients are covered in scenario 1. Additionally, it is difficult to get people to travel to a central location for a screening day because financial constraints or daily routines of herding cattle or taking care of the homesteads often restrict travel. Community testing, as depicted in scenario 4, is dependent on a multiplexed POC test that tests for multiple diseases simultaneously. Multiplexed tests are often more expensive and very few, if any, have been routinely used in the field.

Similarly, integration with the malaria journey, as depicted in scenario 5, is also dependent on a multiplexed POC test for VL and malaria. A well-known multiplex test from DIAMED for VL and malaria is relatively expensive and not routinely used in Kenya and Uganda. By contrast, malaria RDTs are affordable and are routinely used in the field. A follow-up test, as depicted in scenario 6, would be challenging to implement due to the nomadic nature of the inhabitants that we encountered.
A unidirectional problem-solving approach, instead of an interactive design-thinking approach that requires multiple iterations, would have yielded possibly 1 or 2 predictable scenarios instead of the detailed thorough 8 scenarios that were obtained in this study. Thus, after multiple iterations with a Kenyan public health officer and Médecins Sans Frontières, we concluded that scenarios 1 and 2, a test for screening and confirmation and a test-of-cure, are pivotal for VL management as they meet the local need and are feasible to implement. Furthermore, the CTPP developed for the 2 selected scenarios could be broadly applicable to the other 4 scenarios.

7. Define the CTPP

Within each of these 2 selected scenarios, we identified variables that clarify how the diagnostic setting influences the features of the diagnostic test. A CTPP such as formulated in Figure 8 presents the key features of a product that would fit a particular local health care context. The key features

FIGURE 8. (A) CTPP for a VL Point-of-Care Diagnostic Screening-and-Confirmation Test, (B) CTPP for a VL Point-of-Care Diagnostic Test of Cure

<table>
<thead>
<tr>
<th>Feature</th>
<th>SCREENING &amp; CONFIRMING</th>
<th>TEST OF CURE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why (scope)</td>
<td>To screen key populations and confirm the disease without subsequent re-testing</td>
<td>To determine efficacy of the treatment administered on the identified VL patients</td>
</tr>
<tr>
<td></td>
<td>Preferably a non-quantitative test i.e. providing a positive or a negative result</td>
<td>Preparing a positive or a negative result</td>
</tr>
<tr>
<td>Where (level of the healthcare setting)</td>
<td>At the homesteads or at the mobile health posts - lowest level</td>
<td>At the higher levels of the healthcare systems, primary healthcare clinics and hospitals</td>
</tr>
<tr>
<td>Who (target user of the test)</td>
<td>Healthcare workers (CHW/CHV) with minimum training</td>
<td>Trained and certified laboratory technician</td>
</tr>
<tr>
<td>When (diagnosis moment)</td>
<td>Initial diagnosis i.e. pre-treatment</td>
<td>Post-treatment (relapse of the disease)</td>
</tr>
<tr>
<td>How (operational characteristics)</td>
<td>Sample type: non-invasive (urine) or minimal invasive (finger pin-prick of blood)</td>
<td>Sample preparation with limited number of steps</td>
</tr>
<tr>
<td></td>
<td>Sample preparation should be either fully integrated or with minimum number of steps required minimal user interaction</td>
<td>Extensive user training required</td>
</tr>
<tr>
<td></td>
<td>Minimal user training required</td>
<td>No strict requirements - minimum</td>
</tr>
<tr>
<td></td>
<td>Preferably hand-held device or a small, portable battery operable instrument (&lt;1 kg)</td>
<td>Test should be single-use, disposable (preferably biodegradable or recyclable) with no maintenance or minimal preventative maintenance, if necessary</td>
</tr>
<tr>
<td></td>
<td>Should be able to withstand transportation stress</td>
<td>Safety: closed, self-contained system</td>
</tr>
<tr>
<td></td>
<td>Test should be single-use, disposable (preferably biodegradable or recyclable) with no maintenance or minimal preventative maintenance, if necessary</td>
<td>Operating temperature (+20°C to +50°C) requiring no cold chain for storage</td>
</tr>
<tr>
<td></td>
<td>Safety: closed, self-contained system</td>
<td>Time to result: preferably &lt;30 min and not more than 3 hours</td>
</tr>
<tr>
<td></td>
<td>Operating temperature (+20°C to +50°C) requiring no cold chain for storage</td>
<td>Result display/interpretation: results visible to a naked-eye with minimal instructions for user-dependent interpretation or with an integrated reader with a display - YES/NO/INVALID</td>
</tr>
</tbody>
</table>

Abbreviations: CHV, community health volunteer; CHW, community health worker; CTPP, concept target product profile; VL, visceral leishmaniasis.
that a CTPP takes into account include variables such as why a test is performed (scope); where the test will take place (geographical location); which level of the health care system (such as at a home or a hospital); who the end user is (target user of the test); when the test will take place in the patients’ health-seeking pathway (diagnostic moment); and how the test will be conducted (operational characteristics of the test). In this study, the product is a new POC diagnostic test for VL in rural Kenya and Uganda. A CTPP is an efficient mapping tool that can guide the R&D of a new POC diagnostic test by determining the local needs, the end user(s), the intended use, and the adequate features of the test. We conclude that the approach presented in this study led to the development of a CTPP wherein the essential features of a POC diagnostic test for VL in Kenya and Uganda were identified more quickly compared with a TPP. The added value of a CTPP is that it does not impose overly stringent guidelines on the researcher during the early stages of R&D. Instead, a CTPP aims to guide the research and not limit the potential of the new POC diagnostic test that is under development.

**DISCUSSION**

The global diagnostic need for NTDs demands innovative solutions that are customized to the local health care context. To develop a technology that can be implemented in the relevant context, researchers must get access to a comprehensive, yet easily accessible overview of the status quo to understand the challenges and limitations of the existing health care system. Despite extensive literature reviews on VL diagnostics, treatment, and management, an evident knowledge gap persists, which prompted us to conduct field research in VL endemic regions and develop the notion of a CTPP.

Although it is imperative to conduct a comprehensive context analysis before and during R&D of the technology, this is often done only at a much later stage by specialized teams, which may not include scientific researchers. The present study was conducted by scientific researchers who are developing a POC diagnostic test in collaboration with industrial-design experts to define scientifically feasible innovative solutions for VL diagnosis in a resource-limited setting. In contrast to a TPP that comprises a much larger group of experts such as social scientists and policymakers, a CTPP can be formulated effectively in a fast manner with a team that includes technical researchers, health care experts, and industrial designers.

Initially, it appeared challenging to determine how to effectively engage with local stakeholders during a limited time in the field. From the very beginning of our approach, we overcame many logistical and cultural barriers by engaging with local stakeholders. Apart from engaging with medical professionals (doctors, nurses, and laboratory technicians), we spent valuable time with the CHWs/CHVs and observed their crucial role in VL diagnosis in the field. We observed many similarities in the health care systems between Kenya and Uganda, which is unsurprising given the common border and a shared history; therefore, we created a combined Gigamap of their VL health care systems. We observed many challenges that are encountered by the locals: food and water insecurity, which causes malnutrition; remoteness; and a lack of infrastructure and poor health care systems, which collectively adversely affect access to adequate diagnoses and treatment. Furthermore, the lack of resources, both financial and in the availability of health care staff and RDTs for VL, and poor health-seeking behavior, which is exacerbated by the lack of education that is prevalent in the remote regions, impede effective VL management.

Conducting field research provided rich sources of information for understanding how new POC diagnostic tests can fit into a specific health care context. We would like to highlight the importance of international cocreation through active collaborations between all stakeholders, including academia, industry, nonprofit, and governmental organizations.7 Cocreation with local experts is necessary to understand the implementation need for a new POC diagnostic test, as well as to ensure the sustainable use of the POC test in the field by building trust and mutual interest and creating a foundation for knowledge transfer to engage locals in the future. New technologies are often mistrusted by health care providers, especially in remote settings, due to a lack of understanding of the complex research behind the development of the new technologies. Thus, cocreation with the local stakeholders starting from the design phase (conceptualization) to the prototype phase (realization) ensures a strong relationship with the end users. The approach that we adopted, in fact, promoted collaboration (and not competition) between the key stakeholders.

A number of key findings that were learned from the approach are as follows:

- A CTPP is an effective new tool that can aid R&D researchers in matching the technology that they develop to a specific health care...
context more quickly than a conventional target product profile.

- The role of local volunteers and community health care workers is critically important for access to diagnostics in resource-limited settings. With improved yet simplified VL POC diagnostic tests, CHWs/CHVs could perform diagnosis of VL closer to the homesteads.

- A noninvasive test-of-cure and a screening and confirmation test will significantly improve the management of VL in the endemic regions. This would greatly benefit patients, particularly immunocompromised patients who are at a higher risk of relapse, as well as help pharmaceutical researchers and clinicians who are developing and testing new VL treatment regimens.

- The cost of the diagnostic test is an important factor to consider, especially in view of the fact diagnoses needs to be repeated, to screen the patient in the field initially, and again at the treatment facility to rule out any procedural error and to justify the toxicity and expense of VL treatment. Thus, the diagnostic test needs to be affordable.

Key implications learned from the approach are:

- Early during the R&D stage, researchers should consider who will administer the test (patient, health care worker, doctor) and for what purpose.

- Program managers should consider that the training level of staff and volunteers and the availability of the resources are the critical determinants for using a diagnostic test.

- Researchers should consider that introducing themselves to local communities and stakeholders early will improve the willingness of the communities to implement the new technology. Upon further development, testing prototype devices can be facilitated by the local East African partners, thus strengthening international cocreation, and increasing the probability of success of a new POC diagnostic test beyond a mere proof-of-principle.

The plethora of information gathered in the field was comprehensively processed using our methodology, which includes design-thinking tools such as visual thinking, leading to the development of a Gigamap, patient journeys, and the consequent use case scenarios that are presented in this study. The visual thinking was used as a means to summarize, validate, and communicate key insights from the field research to the stakeholders, as well as to create an aligned vision within the team. The visualizations allowed us to identify where, when, and how a new POC diagnostic test can fit into the health care system within a resource-limited endemic region, in the form of a CTPP, which is an efficient mapping tool compared with a TPP. A CTPP approach was applied to sufficiently scope the problem of VL diagnostics, gather contextual information, and define the adequate features of a new POC diagnostic test.

**CONCLUSION**

Disease eradication requires improved diagnostic tests, as well as an efficient system to successfully deliver and implement them in the appropriate settings. As this is largely dependent on the local capacity and the willingness of key stakeholders to participate, solely designing a POC test for a particular setting does not ensure successful implementation of the test. Designing a product for the end user is complex since a wide range of political, social, cultural, and environmental factors contribute, but it is worth the added time, effort, and resources to realize a successful development and implementation of a new POC diagnostic test.

In this article, we presented an approach that included design-thinking principles to formulate a CTPP that consists of multiple steps. Our approach moved from identifying gaps in current VL diagnosis in endemic regions by critically reviewing the existing literature, to selecting an endemic region to validate the literature findings and conduct direct observations in the field. After that, we used visual thinking to create Gigamaps and patient journeys based on the combined insights that were obtained from the literature and the field research, which led to valuable use case scenarios that describe the ideal setting for a new POC diagnostic test. Finally, we used these collective data to formulate a CTPP for a new POC diagnostic test that is specific for VL diagnostics in resource-limited settings.

In summary, we introduced the notion of a CTPP as an effective toolbox to match the development of a POC diagnostic test and the health care context for its application. More generally, we anticipate that a CTPP will be a useful new tool that enables researchers to match the development of new diagnostic tests or medical equipment, and the local health care context in which they will be used. We envision that a CTPP will enable
Matching Point-of-Care Diagnostic Tests to the Local Health Care Context

Global Health: Science and Practice 2020 | Volume 8 | Number 3

researchers to ruminate on the new product and facilitate the iterative design process—and ultimately benefit global health.

Acknowledgments: We acknowledge Jaco van der Torre, Oskar Franch, and Roel Kamerling for discussions and support. H.N. and D. M. acknowledge the National Research Fund from Kenya for funding and support. We thank Dr. Daniel Masiga (International Centre of Insect Physiology and Ecology [ICIP], Kenya), Mr. Johnstone Inganga, Mr. Anyona Joseph (Kenya Medical Research Institute [KEMRI], Kenya), Dr. Patrick Sagaki, Dr. Lawrence Okello Awany, and Dr. Andrew Munyema (Amadat hospital, Uganda) for their immense support, assistance, and contribution to formulate patient stories. We also thank Dr. Koert Ritmeijer and Ms. Nyakia Kamau from Médecins Sans Frontières (Amsterdam) for their expert input and continued support. Lastly, we acknowledge the CHVs/CHWs, academicians (researchers) and medical doctors, technical staff from various health care facilities, and county and subcounty officials across Kenya and Uganda for their time and contribution by sharing their knowledge and experiences.

Funding: Funding was provided by the Delft Global Initiative to conduct the field research.

Competing interests: None declared.

REFERENCES

Factors That Influence Data Use to Improve Health Service Delivery in Low- and Middle-Income Countries

Nicole Rendell, Kamalini Lokuge, Alexander Rosewell, Emma Field

Key Findings
We identified factors that may influence the relationship between information generation and improvement of health services:
- Governance (leadership, participatory monitoring, regular review of data)
- Production of information (presentation of findings, data quality, qualitative data)
- Health information system resources (electronic health management information systems, organizational structure, training)

Key Implications
- Health system researchers should consider how these factors may apply in the field to build a stronger evidence base for how to effectively translate information drawn from health service delivery indicators into improvements in primary health care service delivery.
- Program managers, district level staff, health facility managers, and health care workers should consider what support they need to use available data to improve decision making at the local level and their role in advocating for improved health service delivery in their communities.

ABSTRACT

Background: Health service delivery indicators are designed to reveal how well health services meet a community’s needs. Effective use of the data can enable targeted improvements in health service delivery. We conducted a systematic review to identify the factors that influence the use of health service delivery indicators to improve delivery of primary health care services in low- and middle-income settings.

Methods: We reviewed empirical studies published in 2005 or later that provided evidence on the use of health service delivery data at the primary care level in low- and middle-income countries. We searched Scopus, Medline, the Cochrane Library, and citations of included studies. We also searched the gray literature, using a separate strategy. We extracted information on study design, setting, study population, study objective, key findings, and any identified lessons learned.

Results: Twelve studies met the inclusion criteria. This small number of studies suggests there is insufficient evidence to draw reliable conclusions. However, a content analysis identified the following potentially influential factors, which we classified into 3 categories: governance (leadership, participatory monitoring, regular review of data); production of information (presentation of findings, data quality, qualitative data); and health information system resources (electronic health management information systems, organizational structure, training). Contextual factors and performance-based financing were also each found to have a role; however, discussing these as mediating factors may not be practical in terms of promoting data use.

Conclusion: Scant evidence exists regarding factors that influence the use of health service delivery indicators to improve delivery of primary health care services in low- and middle-income countries. However, the existing evidence highlights some factors that may have a role in improving data use. Further research may benefit from comparing data use factors across different types of program indicators or using our classification as a framework for field experiments.

BACKGROUND
Many countries around the world have developed monitoring and evaluation (M&E) systems to better understand the health of their populations and the effectiveness of their health programs. These systems are intended to capture information about health service delivery to inform how well primary health care services respond to the health needs of a country’s population. Typically, this is achieved through a series of health
service delivery or performance indicators that form part of a broader M&E framework.1-3

The need for M&E has largely been driven by the need for an accountability mechanism in the health system as well as a renewed emphasis on meeting global reporting requirements due to the advent of the Millennium Development Goals and the Sustainable Development Goals.1 Numerous organizations and national governments have developed guidance documents to support development and implementation of M&E activities, including those published by the World Health Organization (WHO), the United Nations Development Programme, and the World Bank.1,4-6 M&E and its component indicators also play a role in continuous quality improvement, which is grounded by a “data use culture” that promotes the use of evidence to inform decision making.7,8

Health service delivery is the operational end point of the health care system, encompassing the provision of a range of services to promote health in individuals that ultimately lead to positive health outcomes in populations.9,10 Health service delivery indicators are designed to leverage the information obtained through routine data collection to gain greater insights into health services and their capacity to meet the needs of the community. Findings from health service delivery indicators can then be used to drive targeted improvements in health services.1,2 However, the success of this process depends in part on how effectively the indicators are used to generate action where change is needed.

The practice of measuring health system performance against a series of context-specific indicators has long been established. The concept of leveraging data or findings from analyses of a set of indicators to improve health system performance (collectively referred to as “data use” or “data-driven quality improvement”) has drawn some attention in the literature.11-15 The evidence on ways in which data use can be enhanced in practice has mostly focused on vertical programs such as immunization and HIV programs,15-18 rather than having used a horizontal system perspective. However, one broad framework that has received attention in the literature is the Performance of Routine Information System Management (PRISM) framework, which incorporates the concept of data use into its assessment tools.19 The framework groups determinants of routine health information system (RHIS) performance into 3 categories—technical, behavioral, and organizational factors. The PRISM framework was developed as a theoretical approach, which has since been tested and validated in a range of settings.19-23 To our knowledge, there has been no comprehensive review of practical strategies that can be employed to promote data use at the primary care level of the health system, across all services.

The purpose of our systematic review was to analyze the current literature to identify what factors influence the use of health service delivery indicators to improve delivery of primary health care services in low- and middle-income countries. Specifically, we focused on the factors that serve as barriers or enablers for national and subnational health authorities to use health service delivery indicators in taking action to improve delivery of primary health care services in low- and middle-income countries.

## METHODS

This systematic review was conducted in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.24

**Health Service Delivery Indicators**

Health service delivery is often described by referencing the type of care (e.g., health promotion, disease prevention, treatment, rehabilitation, palliative care) or the context of the care setting (e.g., ambulatory, primary care, in-patient care).9 For the purposes of our review, a health service delivery indicator is a type of health system performance indicator that is produced routinely and focuses on the operational end point of the health system. To answer the research question, we adopted a broad interpretation of “use of health service delivery indicators” by incorporating key terms into our search strategy to reflect the concept of routinely collected health services data such as service delivery indicators, performance indicators, implementation of M&E systems, and data use. At its core, the research question is about the relationship between information and its impact on health care delivery, and broadening the scope of the search strategy captures the wide variation of terminology in the literature.

**Search Strategy and Selection Criteria**

We systematically searched 3 databases (Scopus, Medline, and the Cochrane Database of Systematic Reviews) and undertook a Google Advanced Search (first 300 citations, using privacy mode) on March 21, 2020. These databases were selected because they are highly regarded in the field of
Factors That Influence Data Use to Improve Health Service Delivery

We manually reviewed reference lists of all included studies were also searched for further eligible studies.

Studies from both the peer-reviewed and gray literature were eligible if they contained empirical evidence on the use of routinely collected health services data in assessing health care quality at the primary care level. The inclusion criteria also required that studies be available in full text and in English. Studies that were set in high- or upper-middle income countries, according to the World Bank Country and Lending Groups, at the time the study was published were excluded. Studies that did not contain empirical evidence and only canvassed a theoretical discussion of the use of routinely collected health services data were excluded.

Two reviewers independently screened all titles, abstracts, and full-text articles according to the inclusion and exclusion criteria. Differences were resolved by consensus. Where findings of a particular study were reported across 2 papers, only the most comprehensive was selected for full-text review.

Data Extraction

For each included study, we extracted information on the study design including setting, study population, and study objective. Key findings relating to the use of routinely collected health services data were also extracted, as well as any identified lessons learned that had potential to provide insights into the research question. Only those findings and lessons learned that addressed the research question by describing factors that influenced the use of routinely collected health services data were extracted and synthesized for analysis.

Quality Assessment

We used the Mixed Methods Appraisal Tool (MMAT) (v-2018, McGill University, Montreal Canada) to undertake a quality assessment. The MMAT was chosen as the framework to assess quality due to its flexibility in assessing different types of empirical studies. Two reviewers independently appraised each of the included studies. We adopted a similar approach to Burnett et al (2018) and classified studies according to the following:

- High quality, if more than 90% of the relevant criteria were met
- Medium quality, if between 60% and 90% of the relevant criteria were met
- Low quality, if between 0% and 60% of the relevant criteria were met

The search terms used were constructed into 3 syntaxes to capture empirical evidence at the primary care level of the health system: (1) "[service delivery indica*" OR "performance indica*"] AND ["health care" OR "health system"] AND [community OR "primary health care" OR "primary care" OR decentrali*ed OR "periph* health cent*"]; (2) "[monitoring and evalu*" AND "implement*"] AND ["health care" OR "health system"] AND [community OR "primary health care" OR "primary care" OR decentrali*ed OR "periph* health cent*"]; and (3) "data*" AND ["information culture" OR "information management" OR "decision?making"] AND ["health care" OR "health system"].

A separate strategy was developed to search the gray literature to capture reports published online by national governments or nongovernmental organizations. Such reports were considered likely to contain valuable insights in the use of health service delivery indicators, but they would not have been identified through searching academic databases alone. The strategy for searching the gray literature using Google Advanced Search was adapted from another systematic review by Graham et al.25 The search terms for “all these words” were data use, monitoring, evaluation, indicators, and performance. These search terms were combined with the following “exact phrases”: health service delivery, primary health care, and primary care. We also directly searched websites of organizations associated with development assistance including Measure Evaluation; UK Department for International Development; German Office for International Development (GIZ); European Commission–International Cooperation and Development; Japan International Cooperation Agency; and the WHO Alliance for Health Policy and Systems Research. Citations identified through the gray literature search were required to meet the same inclusion and exclusion criteria as those from the databases.

All searches were limited to publication dates between 2005 and present. This timeframe was selected because the first decade of the 2000s marked the beginning of sustained momentum in the development of health information systems and health care quality indicators globally.26–28

We manually reviewed reference lists of the systematic reviews within the field of health systems research (rather than those assessing direct interventions) that were set in low- and middle-income countries, to identify eligible studies for inclusion in the full-text review. The reference lists of all included studies were also searched for further eligible studies.
Low quality, if between 30% and less than 60% of the relevant criteria were met

Very low quality, if less than 30% of the relevant criteria were met

We did not undertake an individual assessment of bias for each of the included studies because of the qualitative design of our review.

**Data Analysis**
Themes were derived by undertaking a content analysis of extracted data. Identified themes were considered an enabler if they supported, facilitated, or improved the use of routinely collected health services data to take action to improve service delivery. They were considered a barrier if they restricted, constrained, or prevented this process.

**RESULTS**
We identified 7,393 articles through the peer-reviewed literature search and an additional 289 records from other sources. After the removal of duplicates, 7,340 articles remained and were screened based on their title and abstract. Of these articles, 7,321 were excluded. We assessed 19 full-text articles for eligibility, of which 7 articles were excluded because there were no outcomes relating to use of health services data (n=6) or the study design presented only a theoretical discussion (n=1) (Figure). We identified a total of 12 records from database searches and other sources that met the inclusion and exclusion criteria. No unpublished or in-process studies were identified. An additional 6 reports meeting the inclusion criteria were identified through the manual gray literature search; however, because all 6 reports were from the same source...
Factors That Influence Data Use to Improve Health Service Delivery

(See Supplement for a summary of findings).

**Study Designs**

The majority of the studies that we identified had a cross-sectional design and were conducted in a range of settings. Our sample included a study based in a subnational area of Nigeria; a study based in a national level in Afghanistan that analyzed changes in health system performance over a 5-year period; a study based in 3 districts of Cambodia reporting on implementation of an initiative centered around performance-based financing; a study based in a subnational area of India that evaluated a well-established health management information system; and an evaluation based in Côte d’Ivoire assessing change in quality, availability, and use of data following an intervention that aimed to improve RHIS performance. In addition, there was 1 randomized controlled trial set in Uganda that examined the effectiveness of a community monitoring intervention by comparing communities that received the intervention with communities that did not. The remaining identified studies were either case studies or qualitative studies that were set in Afghanistan, Cambodia, Kenya, Mozambique, Zambia, Rwanda, and Uganda (see Supplement for a summary of findings).

A range of recurring themes emerged from the included studies on data use in practice at the primary care level and the associated factors that enhanced the response to findings from health service delivery, performance indicators, or from M&E activities. Each of these is described in Table 1.

A manual search of the gray literature found 6 reports from Measure Evaluation that met the inclusion criteria. Given that these reports were from a single organization, we present the findings from these papers separately in a secondary gray literature analysis.

**Quality Assessment**

According to the MMAT quality assessment, most of the studies selected for inclusion in the systematic review were medium quality (n=6), with only 2 being assessed as high quality. The studies identified through the review occupy the lower levels of the evidence hierarchy. Only 1 study had a specified control group and an intervention group as part of the design (a randomized controlled trial), which is the most robust design for comparing strategies that enhance data use.

**Analysis—Peer-Reviewed Literature**

**Leadership**

Leadership and the role of active engagement from senior management was highlighted across multiple included studies as a feature associated with strengthened M&E capacity and facilitating uptake at the local level. The concept of leadership is itself nuanced and difficult to measure and is represented slightly differently in each of the studies. In the study by Holvoet and Inberg, which compared Rwandan and Ugandan health systems, there was evidence demonstrating the role of leadership in both countries. In Rwanda, strong leadership was identified as a contributing factor in situations in which evidence was used to effectively remedy an issue. Effective governance and strong linkages between processes for planning and M&E were also found to be important factors. In Uganda, a biannual meeting with ministers and permanent secretaries, in which health sector performance is reviewed and discussed, was shown to improve interest in data quality and use. Similarly, in another study that was also based in Uganda, Kanana et al observed that involvement of management at a local level (health district leaders, health facility managers, and subcountry leadership team) in planning and M&E processes strengthened managers’ capacity to use available data to advocate for change. In Cambodia, Khim et al compared service delivery across 3 districts during the same time period and attributed the success of the highest performing district to strong leadership and management capacity within the district (managers were perceived to objectively undertake performance monitoring). While these 3 studies present insights into leadership and data use, their quality was determined to be low or very low. An earlier study by Jacobs et al, which was also set in Cambodia, found that improvements in aggregated performance by all health facilities occurred when the district health technical advisory team became more actively involved, highlighting the role of active engagement from senior management. Edward et al compared service delivery over a 5-year period across most districts in Afghanistan and likewise observed that leadership, specifically the use of “champions,” was an important factor in the successful uptake of a Balanced Scorecard (BSC) as a performance management tool. These studies were of medium and high quality, respectively. Overall, these
findings suggest that engaged leadership serves as an enabler for using data on health service delivery, in the form of indicators or M&E findings, to drive action in health service delivery improvement.

**Participatory Monitoring**

Other studies explored the concept of community or participatory monitoring to varying degrees.\(^3\)\(^6\),\(^3\)\(^7\),\(^3\)\(^9\),\(^4\)\(^1\),\(^4\)\(^3\) A community monitoring intervention formed the cornerstone of Björkman and Svensson’s study\(^3\)\(^7\) based in Uganda. They found that community monitoring, through dissemination of a report card followed by a series of meetings and joint action planning between the community and health facilities, led to improvements in service utilization and health outcomes. This medium-quality study was the only randomized study with an intervention and control group selected for inclusion in our systematic review. In another medium-quality study, which was based in Côte d’Ivoire, Nutley et al\(^4\)\(^2\) evaluated the impact of a comprehensive data

<table>
<thead>
<tr>
<th>Reference and Location</th>
<th>Peer-Reviewed Literature Analysis</th>
<th>Gray Literature Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Peer-reviewed literature analysis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Björkman and Svensson(^3)(^7) – Uganda</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Chukwuani et al(^3)(^2) – Enugu State, Nigeria</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Edward et al(^3)(^3) – Afghanistan</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Edward et al(^3)(^9) – Afghanistan</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Holvoet and Inberg(^3)(^8) – Rwanda and Uganda</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Jacobs et al(^3)(^5) – Kirivong Operational Health District, south east Cambodia</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Kananura et al(^3)(^1) – 3 districts in eastern Uganda</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Kim et al(^3)(^4) – Cambodia</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Krishnan et al(^3)(^5) – Ballabgarh, India</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Nutley et al(^4)(^2) – Kenya</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Nutley et al(^3)(^6) – Côte d’Ivoire</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Wagenaar et al(^3)(^3) – Mozambique, Rwanda, and Zambia</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Gray literature analysis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Afe et al(^4)(^4) – Nigeria</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Anasel et al(^4)(^6) – Tanzania</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Li et al(^4)(^5) – Tanzania</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>MEASURE Evaluation(^4)(^7) – Mali</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEASURE Evaluation(^4)(^8) – Kenya</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Millar et al(^4)(^9) – Kenya</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Abbreviations: HMIS, Health Management Information System; PBF, performance-based financing.

\(^a\) Identified as potential factors following gray literature analysis.

Findings suggest that engaged leadership serves as an enabler for using data to improve health service delivery.
use intervention. They used dichotomous indicators to assess data use and found it had improved over a 4-year period. The successful intervention incorporated different platforms (e.g., quarterly forums and working groups) to engage a range of stakeholders from data producers to data users. A high-quality study by Edward et al also found that participatory monitoring led to positive outcomes, including increased ownership and accountability among those engaged in the process, improved community awareness of rights in accessing health services, and improved service utilization. However, the authors emphasized the important role of facilitation in meetings between community members and health providers to balance their respective demands. A similar finding was observed by Kananura et al, who found improved engagement from the community and health services management staff in response to participatory monitoring processes designed to monitor implementation of an antenatal care project. Yet the process was only maintained in the presence of the project team, suggesting challenges exist in ensuring long-term sustainability. Further, the study by Wagenaar et al, which was conducted at a system level across 3 countries, observed that shared responsibility of data interpretation and collaborative performance review promoted a culture of data use. Both the Kananura and Wagenaar studies were of low quality. All of these findings indicate that participatory (or community) monitoring may serve as an enabler in facilitating the use of health service delivery indicators.

**Presentation of Findings**

More specifically, in terms of presentation of indicator findings, the 2 studies from Afghanistan and 1 study from Kenya provided evidence that data visualization can be used effectively as a performance management tool when combined with leadership and community participation. The earlier study by Edward et al reviewed performance trends in 28 of 35 provinces over a 5-year period using a BSC, which served as a means of integrating key performance indicators with benchmarks that align with strategic goals in a range of domains. It was found to be a successful way of assessing and improving health service delivery, although a need for design changes to ensure continued relevance over time was noted. A later study by Edward et al examined the feasibility of a variation of the BSC, the Community Scorecard (CSC). A CSC is similar to the BSC, but it seeks community perceptions as part of the analysis and targets the local health care context. Unlike the earlier study by Edward et al, which was set in 28 provinces, the study by Edward et al was set in 3 provinces. The authors found that the CSC has potential as a mechanism for enhancing social accountability for quality of care in primary health care facilities, although the CSC development process required strong facilitation by the research team. Both studies were determined to be high quality. In addition, another type of data visualization, the District Health Profile (DHP) tool, was examined by Nutley et al in Kenya and found to be effective in facilitating decision making at the district level. The authors suggested that the factors leading to its successful implementation included the tool’s focus on programmatic questions (rather than a long list of indicators) to meet specific information needs of the district and the use of existing technology. Commonly cited barriers among their respondents were a lack of computers and other office equipment such as printers and an underlying lack of value placed on data. This study was found to be of medium quality. These findings show that the presentation of findings can enhance health service responsiveness to data, although measures need to be taken to ensure ongoing relevance to the given context and availability of appropriate tools. As such, the presentation of findings may serve as a barrier or an enabler depending on the process used to develop the design.

**Data Quality**

The reliability of routinely collected health services data may affect its use according to 4 of the papers included in our systematic review. A low-quality study by Holvoet and Inberg identified poor data quality as a reason for low levels of data use in Uganda. This study described a self-perpetuating cycle in which limited use of data affected the motivation of health facility staff to improve data quality, which then further reinforced the low use. In contrast, Nutley et al found that implementation of a data visualization tool in Kenya resulted in improvements in data quality even though the primary purpose of the tool was to improve program monitoring to make informed service delivery decisions. This outcome was attributed to the tool helping users identify discrepancies in the data, which could then be corrected. In a separate study by Nutley et al, a data use intervention in Côte d’Ivoire that
included PRISM assessments and periodic data quality audits was found to be successful in improving data use. Both studies were medium quality. Similarly, Wagenaar et al.43 advocated for the introduction of data quality assessments as part of any intervention designed to improve data use. They explained that such assessments promote data use in 2 ways: by ensuring confidence in the data, and by showcasing that change is possible through collective effort.43 However, this was found to be a low-quality study. While improving data quality is typically perceived as a goal in itself, these findings suggest that data quality is linked to data use. They also suggest that high-quality data may serve as an enabler to facilitate use of health service delivery indicators.

Qualitative Data
Qualitative data played a role in supporting quantitative findings in 2 of the included studies.32,41 A study by Chukwuani et al.32 in Nigeria assessed primary health care operations, using a mix of data collection methods. Qualitative audits of primary health care facilities were found to be valuable in uncovering operational problems because respondents were more pragmatic than in the questionnaire about their needs. For example, results from the staff questionnaire suggested that respondents had knowledge of operational plans and the activity schedule, yet the qualitative audit revealed that their knowledge was limited to immunization activities. Interestingly, the community sample did not have a similar difference, and the authors proposed that a community questionnaire alone could provide sufficient information on its perspective of primary health care operations.32 The authors concluded that assessment of primary health care using quantitative data provides valuable information for planning, while qualitative data provide valuable information for understanding effective operations management.32 This study was found to be of medium quality. A more recent study by Kananura et al.41 which supported these findings, observed that using both qualitative and quantitative data and discussing the results with a diverse group of stakeholders allowed deeper exploration into unanticipated or complex issues. However, this study was found to be of very low quality. These findings suggest that qualitative data represent an enabler for using data to improve health service delivery, when they are part of a broader data collection strategy.

Electronic Health Management Information System
The use of an electronic health management information system (HMIS) was captured by 2 of the included studies.35,36 An evaluation of the electronic HMIS was undertaken in Ballabgarh, India, by Krishnan et al.35 The authors found that health workers perceived the electronic HMIS as a useful, time-saving means to improve service delivery through development of a monthly work plan based on available data. The authors also found that program managers perceived the electronic HMIS as a better tool for monitoring, supervision, and data management.35 These findings are consistent with those of Nutley et al.36 who evaluated a data use intervention in Côte d’Ivoire that included implementation of monthly reports from an electronic HMIS at the facility level of the health system.36 Both studies were found to be of medium quality. While an electronic HMIS may have the potential to serve as an enabler, in isolation it may be considered as a necessary but insufficient way to enhance the use of health service delivery data to promote health service delivery improvements.

Organizational Structure
Few of the included studies identified staffing arrangements as a strategy to improve data use.36,38,43 The study by Holvoet and Inberg38 referred to a specific position at the local level to support M&E activities. These authors identified that appointment and training of data managers in health centers in Rwanda strengthened the local M&E capacity.36 However, this study was found to be of low quality. The data use intervention found to be successful by Nutley et al.36 also included M&E-specific positions, and it was determined to be a medium-quality study. The intervention also included additional support for staff such as a leadership program and development of supervision guidelines and data management manuals.36 Wagenaar et al.43 proposed a different perspective by stating that data use interventions should focus on system-wide activities, such as mentoring and supervision and action-planning across all health system actors, rather than on individuals. This study was low quality. These findings point to skilled staff as a possible enabler in supporting the use of health services data to drive change in delivery of health services, although their success may be context specific.

Contextual Factors
Two of the included studies noted that contextual factors such as local politics and available re-
The major themes identified across the gray literature reports included leadership, electronic HMIS, regular reviews of the data, and training in data use.

Performance-Based Financing
Performance-based financing (PBF) is a clear application of health service delivery or performance indicators designed to provide improvements in service delivery through financial incentives for health care providers. The implementation of PBF intersects with M&E activities. The study by Jacobs et al.40 in Cambodia found that the use of performance management in the form of PBF contributed to maintaining a consistent level of health service delivery during a major period of transition. They also found that effective implementation was associated with leadership and was contingent on the M&E activities being undertaken by an independent body.40 Furthermore, Holvoet and Inberg38 found that PBF reinforced the value of data use at the local level. These findings suggest that PBF could be an enabling factor in the use of data for improvement in health service delivery. However, it would not be implemented as a strategy to improve data use in practice, so it is not useful to consider it a mediating factor in the context of our research question.

Analysis—Gray Literature
Papers identified through the manual gray literature search were all from the organization Measure Evaluation. These papers formed the basis of a secondary analysis that was conducted separately from the primary analysis of peer-reviewed literature to minimize the risk of bias. If included in the primary analysis, these papers would have constituted a third of the studies and could have influenced the results. In addition, each paper represents an evaluation of a Measure Evaluation project, by Measure Evaluation. While each was subjected to MMAT quality appraisal, this framework does not accommodate questions of independence of the evaluation.

Six reports across 4 countries—Nigeria, Tanzania, Mali, and Kenya—met the inclusion criteria.44–49 According to the MMAT quality assessment, most of the studies selected for inclusion in the gray literature analysis were of high quality (n=3), with only 1 being assessed as low quality.

The major themes identified across the gray literature reports included leadership,44–48 electronic HMIS,44–46,47,49 regular reviews of the data,45,48,49 and training in data use.44–46,48 This adds weight to the evidence for the role of leadership and electronic HMISs identified in the primary analysis and introduces 2 new potential themes—regular review and training. The gray literature reports also provide evidence to support participatory monitoring,46,48 presentation of findings,45,49 and data quality44,46 as factors in promoting data use. In contrast, the gray literature reports do not provide evidence to support the themes of qualitative data or organizational structure; however, this does not diminish the value of these established themes.

Regular Review
Regular review of program data emerged as a possible factor from the gray literature analysis, and it was subsequently identified as a theme across papers in both analyses. In the analysis of peer-reviewed studies, regular review as a mediating factor in data use was interdependent on other factors, including participatory monitoring or presentation of findings. Consequently, regular review was overshadowed as a standalone theme. Regular data reviews outlined in the gray literature analysis included periodic meetings specifically to understand the data and discuss program performance,45,48,49 technical working groups,48 and stakeholder forums.48 These findings from the gray literature analysis suggest regular data review is an underlying factor, and perhaps a necessary or sufficient condition, that may potentially contribute to the use of health service delivery indicators.

Training
The role of training and capacity-building activities in data use emerged as an independent theme from the gray literature analysis. Although it did not feature strongly in the primary analysis, it is broadly linked to the theme organizational structure, which was prominent. The evidence suggests that the relationship between data use and training is straightforward—training in data use facilitated staff use of data at the local level,36,44,45 and conversely, an absence of training was cited as a barrier to data use.46 The analysis by Measure Evaluation48 in Kenya found that capacity-building activities, which may be considered an extension of training by incorporating ongoing technical assistance and mentoring, resulted in an increased appreciation and ownership of data.
being used in decision making. This finding suggests that capacity-building activities may have a broader reach, facilitating cultural change rather than just an advancement of the technical skillset. Regardless, investment in staff professional development in data use may be an enabler for using health service delivery indicators.

Classification of Influential Factors on Data Use
The factors identified from our analyses may be grouped into categories to facilitate further discussion and research. We propose 3 groupings: governance, production of information, and health information system resources (Table 2). PBF and contextual factors have been excluded from the classification because they do not represent mediating factors that can be adapted in the short to medium term in practice.

DISCUSSION
We identified 12 published studies and 6 reports from a range of low- and middle-income countries that provided empirical evidence on factors that influence the process of using health service delivery indicators to improve delivery of primary health care services. The low number of studies meeting our inclusion criteria suggests that this area of research has received little attention, making it difficult to draw reliable conclusions. Most of the influential factors identified in this setting appeared to serve as enablers. These included the role of leadership in facilitating the use of indicator findings at a local level, participatory (or community) monitoring, presentation of findings, data quality, qualitative data, electronic HMIS, and organizational structure. Regular review of data and training in data use may also have roles to play as independent factors, however, supporting evidence is less clear. The influential factors were grouped into 3 categories for further discussion: governance, production of information, and health information system resources (Table 2).

Contextual factors and PBF were each found to have a unique relationship with the use of health service delivery indicators, but they may not be practical to discuss as mediating factors in terms of promoting data use.

Governance
The studies in our systematic review that discussed the role of leadership referred to engagement of health service managers and senior executives in the planning and evaluation processes at the local level. The importance of leadership in the context of effective health services management is widely acknowledged, so it is not surprising that strong leadership may be an enabling factor in promoting use of service delivery indicators. However, the challenge lies in understanding the ways in which leadership capability can be strengthened. Much of the existing literature on leadership in the health sector is focused on individuals. There is scope to undertake more multi-level analyses that consider different team and organizational factors. One report by the Alliance for Health Policy and Systems Research examined participatory leadership as a strategy for improving health systems. The report proposed that participatory leadership draws on the collective strength of different actors across the health system that can have a stabilizing impact (reduces vulnerability to actions of individual leaders) or a disruptive impact (challenges the status quo as needed). This concept aligns with the studies identified in our review that highlight multilevel engagement as a form of strong leadership.

The terms community monitoring and participatory monitoring are used interchangeably and are both forms of social accountability. Social accountability is described by Hamal et al. as “the mechanisms that citizens can use to hold the state and service providers to account for their actions.” Some evidence supports the use of social accountability mechanisms, such as participatory monitoring, to promote quality improvement, particularly

<table>
<thead>
<tr>
<th>TABLE 2. Classification of Influential Factors on Data Use to Improve Health Service Delivery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Governance</td>
</tr>
<tr>
<td>• Leadership</td>
</tr>
<tr>
<td>• Participatory monitoring</td>
</tr>
<tr>
<td>• Regular review&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup>Factors identified through gray literature analysis.
In the studies identified by our review, data quality could be credited with contributing to improvement in the use of health service delivery data.

**Production of Information**

Our review identified 3 studies that examined 3 different platforms for presenting the findings drawn from monitoring data and service delivery indicators, the BSC, the CSC, and the DHP tool. The literature around BSC and CSC scorecards suggests that CSC is an extension of the BSC, which engages the community and is also perceived as a social accountability mechanism (see previous section on participatory monitoring). Unlike the DHP tool, which was developed as a national solution to support district health data integration in Kenya, there is evidence to support effective use of both the BSC and CSC in other settings, outside Afghanistan, including high-income countries. With the advent of electronic HMISs, the opportunities to develop data visualizations such as scorecards and dashboards have grown as has the capacity to measure system logins and data use by health staff. A recent “realist” review of immunization data use undertaken by PATH found moderate-certainty evidence that decision support tools such as dashboards may improve data use. It also found that such tools are most effective when integrated with established data review and decision-making processes and other forms of feedback such as supportive supervision.

In the studies identified by our review, data quality could be credited with contributing to improvement in the use of health service delivery data, at least in part. While each study proposed an explanation to account for the relationship, it is unclear if the link is due to staff motivation, reassurances in data accuracy, or another factor related to the specific setting of our research question. The realist review by PATH, which focused on use of immunization data, indicated poor quality may be a barrier to using data, but better data quality does not lead to improved use. However, there is evidence that suggests the reverse: improved use of data may improve data quality.

Numerous guidance materials, such as those published by WHO, the World Bank, and the United Nations Development Programme, have advocated for the use of qualitative data as part of a mixed methods design for M&E activities. The literature also contains many examples of qualitative data being used as part of mixed methods design for M&E activities. The rationale for using qualitative data is to improve data validity, reliability, and credibility.

**Health Information System Resources**

Health information systems form one of WHO’s building blocks and are considered a key part of governance functions. Electronic HMISs are an efficient tool to serve this function and offer numerous advantages to other systems by ensuring data is quality checked at entry; however, they alone do not ensure data quality and use. Our review found mixed results in terms of organizational structures to facilitate use of health service delivery data, with 2 studies proposing specific M&E positions and 1 study advocating for investments in system support structures rather than individuals. In terms of improving M&E capability at the local level, the literature has tended to focus on health workers and the importance of feedback mechanisms, supportive supervision, and training in data quality and use, rather than investment in additional trained staff. This focus is consistent with our gray literature analysis, which identified training as a potential mediating factor in and of itself.

PBF is a supply-side provider payment mechanism that uses financial incentives to motivate individuals and organizations.
The main limitation of this review concerns the approach.98,99 with local stakeholders, and is in line with the HCD such as Tanzania and Ethiopia using partnerships. The work is funded by the Bill and Melinda Gates Foundation, has been implemented in countries to better decisions and improved health outcomes.97 regular data use will create a data use culture, leading of change model that hypothesizes better data and use culture. These partnerships are based on a theory of change that centers around user needs and applies design thinking principles such as prototyping.96 Recent years have seen the introduction of data use partnerships, which aim to build a sustained data use culture. These partnerships are based on a theory of change model that hypothesizes better data and regular data use will create a data use culture, leading to better decisions and improved health outcomes.99 The work is funded by the Bill and Melinda Gates Foundation, has been implemented in countries such as Tanzania and Ethiopia using partnerships with local stakeholders, and is in line with the HCD approach.98,99

Limitations
The main limitation of this review concerns the search strategy. We hypothesized that potentially valuable information is contained in evaluation reports that are published in the gray literature on websites of government agencies or nongovernmental organizations. However, identifying an appropriate methodology that would capture such reports was challenging. During the design phase of this systematic review, different strategies were tested using gray databases and deep web search engines. However, their results could not be repeated for our setting. As such, we decided to opt for a simplified approach using a standard search engine (Google Advanced Search) in privacy mode. Using this approach, no reports were identified that met the inclusion criteria. We also manually searched websites and identified reports that met the inclusion criteria from a single source. To manage the risk of bias we adopted a 2-tier approach to our analyses so the reports that were not published in the peer-reviewed literature were treated as supplementary information, which limited their capacity to distort the results. The search strategy applied limits to the setting (studies in high-income countries were excluded) and primary care level of the health system. While we acknowledge that some observations in high-income countries could likely be applied to low- and middle-income settings, in the interest of responding to the research question, we chose to adopt a narrow scope. This decision may have excluded some studies that should have been included.

Results may also have been limited by the naming conventions of health indicators. For example, indicators reporting on an immunization program may be recorded in the database as immunization indicators or program indicators even though they would also fit the criteria as a performance or service delivery indicator. Further research looking into the application of service delivery indicators may benefit from investigating only program-specific indicators or selecting some key programs and then comparing and contrasting their use across different settings. Alternatively, our classification could be used as a framework to undertake a series of field experiments similar to the methodology of Björkman and Svensson.37 This approach supports understanding the weight of each factor, its relationship with other factors, and the effectiveness of system-level approaches.

In addition, we did not undertake an objective assessment of bias for each of the selected studies. The broad range of studies selected for this review meant that the MMAT quality assessment tool was selected to accommodate such differences. Although the MMAT did allow for some level of assessment of bias, comprehensively assessing the
risk of bias for each individual study was not possible. The scope of the review may have introduced a publication bias across studies. The research question targets low- and middle-income countries; however, due to limited research capacity in these settings, the literature likely has an underrepresentation of studies from these settings. As such, the experiences of low- and middle-income countries that did not feature in the literature may differ from those published and subsequently captured by the search strategy. The qualitative content analyses may have also introduced an unavoidable risk of both selection and measurement bias.

CONCLUSION
Scant empirical evidence is available on how health service delivery indicators are used to improve primary health care services in low- and middle-income countries. It is clear there is no single known intervention that could be applied in isolation. However, our systematic review identified some factors that may influence the use of service delivery indicators in practice in low- and middle-income settings: governance (leadership, participatory monitoring, regular review of data); production of information (presentation of findings, data quality, qualitative data); and health information system resources (electronic HMIS, organizational structure, and training in data use). Most of these factors are likely to have an enabling effect. Both contextual factors and PBF were found to have a relationship with the application of health service delivery indicators, but it is not useful to consider these as mediating factors in practice.

Given the narrow scope of the search strategy applied in this review, future research may consider undertaking a broader analysis across different types of program indicators and comparing how these drive change. Alternatively, one could use the classification proposed by this review to test interventions associated with each of the factors in the field to better understand the interrelationships and other possible dominant characteristics that promote translation of data into improved health service delivery at the primary care level of the health system.

Acknowledgments: We thank Dr. Colin Wiltshire for his contribution during the conception and design of this systematic review.

Funding: This research is supported by an Australian Government Research Training Program Fee Offset Scholarship.

Conflicting interests: None declared.

REFERENCES
Factors That Influence Data Use to Improve Health Service Delivery


Factors That Influence Data Use to Improve Health Service Delivery

Received: November 12, 2019; Accepted: July 7, 2020; First published online: September 4, 2020


© Rendell et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are properly cited. To view a copy of the license, visit http://creativecommons.org/licenses/by/4.0/. When linking to this article, please use the following permanent link: https://doi.org/10.9745/GHSP-D-19-00388
METHODOLOGY

Mask Reuse in the COVID-19 Pandemic: Creating an Inexpensive and Scalable Ultraviolet System for Filtering Facepiece Respirator Decontamination

Rachel M. Gilbert, a Michael J. Donzanti,a Daniel J. Minahan,a Jasmine Shirazi,a Christine L. Hatem,a Brielle Hayward-Piatkovskyi,b Allyson M. Dang,c Katherine M. Nelson,d Kimberly L. Bothi,a,e Jason P. Gleghorn,a,b

Key Messages

- Ultraviolet germicidal irradiation (UVGI) systems can be used to decontaminate filtering facepiece respirators that are in short supply during the current COVID-19 pandemic, but are costly and scarce.
- Custom-built UVGI systems can be easily and affordably created using common items found in hardware stores and within research institutions.
- Health care workers and administrators should consider this setup as a cost-effective option to combat personal protective equipment (PPE) shortages during the current pandemic.
- Academic institutions should consider fostering collaborations with local health care institutions to provide idle resources to front line health care workers facing PPE shortages.

ABSTRACT

As the current COVID-19 pandemic illustrates, not all hospitals and other patient care facilities are equipped with enough personal protective equipment to meet the demand in a crisis. Health care workers around the world use filtering facepiece respirators to protect themselves and their patients, yet during this global pandemic they are forced to reuse what are intended to be single-use masks. This poses a significant risk to these health care workers along with the people they are trying to protect. Ultraviolet germicidal irradiation (UVGI) has been validated previously as a method to effectively decontaminate these masks between use. However, not all facilities have access to the expensive commercial ultraviolet type C (UV-C) lamp decontamination equipment required for UVGI. UV-C bulbs are sitting idle in biosafety cabinets at universities and research facilities around the world that have been shuttered to slow the spread of COVID-19. These bulbs may also be available in existing medical centers where infectious diseases are commonly treated. We developed a method to modify existing light fixtures or create custom light fixtures that are compatible with new or existing UV-C bulbs. This system is scalable; can be created for less than US$50, on site and at the point of need; and leverages resources that are currently untapped and sitting unused in public and private research facilities during the pandemic. The freely accessible design can be easily modified for use around the world. Health care facilities can obtain this potentially lifesaving UVGI resource with minimal funds by collaborating with research facilities to obtain the UV-C meters and UV-C bulbs if they are unavailable from other sources. Although mask reuse is not ideal, we must do what we can in emergency situations to protect our health care workers responding to the pandemic and the communities they serve.

INTRODUCTION

Health care workers (HCWs) are critical to the care and treatment of individuals with the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) or coronavirus disease known as COVID-19. In addition to the needed beds and ventilators, personal protective equipment (PPE), particularly filtering facepiece respirators (FFRs), are essential to ensure the health and safety
of not only trained doctors, nurses, and emergency response personnel, but also other health care facility staff who play an important role in cleaning, disinfecting, and preparing spaces for patient care. Additionally, whereas much of the focus of HCW risk is on large hospitals and current hotspots that do not have access to FFRs, also called N95 masks in the United States, the spread of SARS-CoV-2 has also affected residential facilities and rural clinics around the globe. These communities face additional challenges with limited resources and larger logistical obstacles to obtain FFRs.

At the time of publication, shipping used FFRs to localized centers for hydrogen peroxide vapor (HPV, also written as H₂O₂ vapor) decontamination was only available in very limited locations in the United States.¹–³ HPV decontamination is a U.S. Food & Drug Administration-approved method for N95 mask decontamination, and manufacturing and deployment of these systems is currently underway. However, the operational and coordination challenges associated with even localized deployment of HPV centers for N95 decontamination are significant. This is evidenced by contemporary reports of HCWs being issued FFRs for continuous (re)use over week-long time periods.⁴–⁶ Even many months after this pandemic entered the global arena, many locations around the world and in the United States are still struggling to provide adequate PPE for HCWs and other staff at high risk.⁷,⁸ Additionally, Nebraska Medicine has initiated a U.S. Centers for Disease Control and Prevention (CDC) approved on-site ultraviolet germicidal irradiation (UVGI) decontamination system for FFR mask decontamination.⁹ UVGI has been demonstrated to be effective at quickly decontaminating FFRs for viruses like the novel SARS-CoV-2 and for multiple cycles of decontamination.¹⁰–¹² Whereas UVGI decontamination has important limitations as discussed herein, these methods currently are being deployed as emergency procedures during the SARS-CoV-2 pandemic. The Nebraska Medicine protocol uses an operating room UVGI decontamination system for FFR mask decontamination, a system that many smaller clinics, rural hospitals, and residential health care facilities may not have.

We document procedures to build a similar type of UVGI platform using off-the-shelf components from a hardware store and UV-C bulbs that can be obtained through online marketplaces or from biosafety cabinets (class I, II, or III) that are ubiquitously found throughout academic research and industrial centers around the world. This system is scalable; can be created for less than US$50, on site and at the point of need; and leverages resources that are currently untapped and sitting unused in public and private research facilities that have shut down during the COVID-19 pandemic. Health care facilities can obtain this potentially lifesaving UVGI resource with minimal funds by collaborating with research facilities to obtain the UV-C meters and limited availability UV-C bulbs required for UVGI treatment.

### FFR Demand and Need During the COVID-19 Pandemic

The COVID-19 pandemic is expected to continue to increase the burden on health care providers. As the number of cases increase, health care facilities will continue to be stretched to their limits in terms of supplies and labor. There are 6,146 hospitals in the United States, 5,198 of which are classified as community hospitals; 209 are federal government hospitals, 616 are nonfederal psychiatric hospitals, and 123 other hospitals.¹³ Community hospitals are those that can be accessed by the general public, including short-term general and specialty hospitals, and are classified as rural or urban. There are 3,377 urban community hospitals that serve approximately 106,000 square miles (about 84% of the population), and there are 1,821 rural community hospitals that serve approximately 3.4 million square miles (about 16% of the U.S. population).¹⁴ These numbers do not include urgent care centers, doctor’s offices, and other non-hospital medical sites involved in the pandemic response.

In contrast, many low- and middle-income countries (LMICs) rely heavily on a limited number of fully resourced hospitals in urban centers with varying degrees of professional health care access in rural areas. Current data on health facilities is difficult to find in many LMICs; however, we can find examples of the resource constraints. For example, Kenya has 842 public and private hospitals serving more than 53 million people.¹⁵,¹⁶ Only 24 of these facilities are classified as county or national referral hospitals and large teaching or private hospitals. The remainder of Kenya’s more than 11,000 health facilities include smaller clinics, dispensaries, health centers, maternity wards, and nursing homes, not to mention thousands of volunteer HCWs in rural communities. Due to limited supplies of PPE at these facilities, the Government of Kenya recently pleaded with the public to reserve N95 masks for the nation’s HCWs because the entire population was required to wear face coverings in public.¹⁷
There are currently dramatic shortfalls in protective equipment in countries with relatively robust health care services, like the United States, and these resources are even more precious in resource-limited communities where fewer doctors and nurses are serving larger populations. According to the World Health Organization18:

Africa suffers more than 22% of the global burden of disease but has access to only 3% of HCWs, and less than 1% of the world’s financial resources.

Losing a single doctor during this pandemic can have a detrimental impact on already strained health care systems across the continent.19 Even in the United States, rural clinics and hospitals serve patient populations sometimes across hundreds of miles, and in some areas, there is a single doctor for several thousand square miles.20 Ensuring HCWs around the world are protected as best as possible is not only ethical, but imperative.

Challenges exist for both urban and rural health care facilities globally. Although urban hospitals often have access to more resources due to the larger population they serve, they experience a strain on their resources during a pandemic precisely because of the significantly larger numbers of people they need to urgently treat. Conversely, rural clinics and health care facilities around the globe often have less funds to operate and face additional logistical challenges to provide patients access to care. Facilities cannot afford to have staff become ill and lead to a decrease in the number of HCWs to treat patients, and therefore need proper FFRs to protect themselves. Additionally, HCWs can potentially spread infection if they are not properly equipped with essential FFRs or if they are forced to reuse potentially contaminated masks. Many HCWs are currently facing the options of not wearing an essential mask, constructing makeshift FFRs with limited efficacy, or reusing a soiled mask—the majority are choosing the latter.5,21,22 There is an unprecedented worldwide shortage of lifesaving equipment that our HCWs need to continue serving their communities safely. In addition, essential personnel working in pharmaceuticals, dentistry, custodial services, delivery services, and law enforcement also require protection while they keep operations afloat. A global shortage of FFRs is expected to persevere due to supply chain challenges, especially for 1 essential component: the melt-blown polypropylene fabric material that filters infectious diseases like the SARS-CoV-2 during inhalation by the wearer.23 As the virus continues to spread into both overburdened and underserved health care systems around the globe, already limited FFR supplies are in extremely short supply. Distributed systems for N95 decontamination are needed to keep up with demand.

The CDC estimated that a 42-day influenza outbreak in the United States, which represents just 4.25% of the global population, could require more than 90 million FFRs for HCWs alone.24 This would scale to almost 800 million FFRs in a year. A model of a hypothetical influenza pandemic predicted 1.7 to 7.3 billion respirators would be required if only 20%-30% of the U.S. population were to be infected.25 This does not account for non-HCWs, such as law enforcement officers and other essential personnel, who may require respiratory protection. Given the uncertain nature of this pandemic and demonstrated logistical challenges in obtaining adequate resources, it is reasonable to assume that need will far exceed the value given in this projection and that demand will only grow.

**FFR DECONTAMINATION AND REUSE**

FFRs are designed and manufactured for single-use applications. Depending on the specific country, FFRs are named differently as masks with the following distinction as N95 (United States), FFP2 (Europe), KN95 (China), P2 (Australia/NZ), Korea 1st class (Korea), or DS (Japan). There is slight variability to the mask filtering specification depending on government regulations, but these masks are all expected to function similarly.26 Each mask is composed of polypropylene fibers to create a physical barrier based on pore size and leverages the electrostatic charge of the material to improve filtering of aerosolized particles such as SARS-CoV-2 viral particles. The inability to scale FFR manufacturing at the rate needed to meet current demand during the COVID-19 pandemic has necessitated the reuse of N95 respirators among HCWs. Work has shown that pathogens such as viruses can contaminate and exist for extended periods of time on the outer surface of FFRs.27 Beyond the risk to HCWs in storing and reusing what is intended to be single-use PPE, other at-risk patients could be exposed to the virus when consulting with a HCW who is reusing their PPE that was previously used with a patient who is positive for COVID-19. It should be noted that repeated redonning FFRs alone poses serious risk to the user due to loss of strap elasticity, nose fit, and therefore, mask integrity. Due to need, redonning is already occurring, and these masks...
have the potential to be contaminated with viral particles, risking further spread of the virus.

The CDC has suggested a method of FFR reuse by issuing each HCW 5 FFRs. Upon completion of a shift, the FFR is placed in a paper bag and redonned after sitting in that bag for 5 days.\textsuperscript{28} This time period was determined based on data that suggests that SARS-CoV-2 can exist on surfaces for as much as or longer than 72 hours. This allows a HCW to cycle through their own previously used masks; however, there are still a variety of problems with this method. Potentially contaminated masks must be stored for long periods of time, and often HCWs do not have the proper space to do this. Masks may also be heavily soiled or moist from the prolonged wear, which would preserve the virus for longer durations, and it has not been rigorously tested whether 5 days is enough time to decontaminate masks under these conditions. HCWs and other essential personnel are desperate for another solution to decontaminate masks effectively, cheaply, and more quickly.

Decontamination of FFRs must be considered carefully because improper decontamination can also give users a false sense of security in addition to compromising mask integrity. A variety of options have recently been developed to allow for decontamination between uses including HPV decontamination, UVGI treatment, and the applications of heat/humidity/washing.\textsuperscript{29} The recently established N95DECON website (https://www.n95decon.org) gives a summary of these methods, including current understandings and limitations to consider for each method. The U.S. Food & Drug Administration approved hydrogen peroxide-based decontamination offered by a U.S. company, Battelle. Although it is a highly effective resource, the procedure relies on shipment of contaminated masks to Ohio, where the company is based, for decontamination\textsuperscript{30,31} and/or the production of such HPV decontamination equipment for regional deployment. The ability for health care facilities and health workers to decontaminate their own masks in minutes, as opposed to days, is a great advantage of using UVGI.

Importantly, UVGI is listed by the CDC as an appropriate method of FFR decontamination and provides key considerations when using this method of decontamination.\textsuperscript{28} Successful implementation of UVGI in a hospital setting is already being used by Nebraska Medicine.\textsuperscript{9} Unfortunately, their system requires 2 surgical suite UVGI towers, with each costing in excess of US$20,000, which not all health care facilities have available.

Herein, we developed a UVGI lamp setup that provides the capability for health care facilities and local regional centers that do not have access to operating room UVGI towers to implement their own N95 mask decontamination system. These health care facilities can use our proposed low-tech UVGI lamp, along with the work flow developed by Nebraska Medicine, to decontaminate FFRs in their own centers. Our UVGI lamp is accessible, inexpensive, requires little expertise to construct and operate, and repurposes existing UV-C bulbs not currently in use. The system takes advantage of common parts available at any hardware store. Once implemented, this method allows for high throughput and quick decontamination cycles that should allow for safer redonning of FFRs.

**FFR DECONTAMINATION USING UVGI**

UVGI systems have been used throughout the health care industry to decontaminate work environments such as surgical suites, equipment, and ambulances. Single-stranded RNA (ssRNA) viruses, like SARS-CoV-2, are especially susceptible to UV decontamination.\textsuperscript{10} Previous work has shown that UVGI systems can also be used to decontaminate FFRs by reducing the viability of the influenza virus, also an ssRNA virus, by 3 log.\textsuperscript{11} A very recent small study has shown that UV is capable of decontaminating N95 mask fabric contaminated with SARS-CoV-2.\textsuperscript{32} There is variable effectiveness of UVGI depending on the mask manufacturer, the different materials of the mask (polypropylene filter versus rubber strap), and the medium in which the virus resides (in liquid, in air, on surface). UVGI decontamination also runs the risk of damaging the FFR materials, which can compromise the integrity of the mask and its usefulness in filtering particles and acting as an effective piece of PPE. However, a variety of studies have looked at the effect of UVGI on mask integrity,\textsuperscript{33,34} even with repeated exposure,\textsuperscript{12,29,35–37} and have found no significant increase in viral penetration, nor decrease in mask stability, even at UVGI doses >10,000 times the required dose to effectively reduce influenza infectivity.\textsuperscript{38} In other studies, 3 UVGI cycles of 1.6–2.0 mW/cm\textsuperscript{2} for 15 minutes did not cause significant changes in respirator fit,\textsuperscript{35} and there was no change in filtration performance. However, using UVGI to decontaminate FFRs requires careful monitoring of UV dosage and the number of times a single mask is decontaminated to minimize damaging the
mask’s integrity. The current literature related to UVGI decontamination of N95 masks, mainly in the context of influenza, is summarized (Table). Importantly, the CDC has approved the protocol from Nebraska Medicine for UVGI decontamination during the COVID-19 pandemic.

UVGI inactivates viruses by damaging their nucleic acids and, to a lesser extent, their proteins.

<table>
<thead>
<tr>
<th>First Author</th>
<th>Year</th>
<th>Dose</th>
<th>Time</th>
<th>Dose</th>
<th>Repeated</th>
<th>Total Accumulated Dose</th>
<th>Distance</th>
<th>Bulb Specs</th>
<th>Effectiveness</th>
<th>Room Conditions</th>
<th>Fit/Degradation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tseng10</td>
<td>2007</td>
<td>5–255 sec</td>
<td>2.51–6.50 mJ/cm²</td>
<td>1×</td>
<td>2.51–6.50 mJ/cm²</td>
<td>30.5 cm</td>
<td>4 × 8 W</td>
<td>Kills 99%, decreased effectiveness at higher humidity</td>
<td>21–28°C, 55% &amp; 85% RH</td>
<td>No observable physical changes, filtration performance not affected, no noticeable changes in airflow resistance</td>
<td></td>
</tr>
<tr>
<td>Viscusi29</td>
<td>2009</td>
<td>0.18–0.20 mW/cm²</td>
<td>15 min/ side</td>
<td>324–360 mJ/cm²</td>
<td>3×</td>
<td>0.97–1.08 J/cm²</td>
<td>40W</td>
<td>21°C, 50% RH</td>
<td>No observable physical changes, no significant change in penetration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bergman36</td>
<td>2010</td>
<td>1.8 mW/cm²</td>
<td>15 min</td>
<td>1.62 J/cm²</td>
<td>3×</td>
<td>4.86 J/cm²</td>
<td>25 cm</td>
<td>40 W</td>
<td>21°C, 50% RH</td>
<td>No observable physical changes, no significant change in penetration</td>
<td></td>
</tr>
<tr>
<td>Heimbuch33</td>
<td>2011</td>
<td>1.6–2.2 mW/cm²</td>
<td>15 min</td>
<td>1.44–1.98 J/cm²</td>
<td>1×</td>
<td>1.44–1.98 J/cm²</td>
<td>25 cm</td>
<td>80 W</td>
<td>≥4 log reduction</td>
<td>Effective decontamination against droplet and aerosolized influenza challenge, no change to fit after decontamination</td>
<td></td>
</tr>
<tr>
<td>Bergman35</td>
<td>2011</td>
<td>1.6–2.0 mW/cm²</td>
<td>15 min</td>
<td>1.44–1.98 J/cm²</td>
<td>3×</td>
<td>4.32–5.94 J/cm²</td>
<td>40 W</td>
<td>&gt;4 log reduction</td>
<td>21°C, 50% RH</td>
<td>Fit not significantly affected after 4 fit tests, no degradation</td>
<td></td>
</tr>
<tr>
<td>Viscusi34</td>
<td>2011</td>
<td>1.6–2.0 mW/cm²</td>
<td>15 min/ side</td>
<td>2.88–3.96 J/cm²</td>
<td>1×</td>
<td>2.88–3.96 J/cm²</td>
<td>40 W</td>
<td>21°C, 50% RH</td>
<td>No changes in fit, odor detection, comfort, or donning difficulty with UVGI– masks were redonned 5×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lore37</td>
<td>2012</td>
<td>1.6–2.2 mW/cm²</td>
<td>15 min</td>
<td>1.44–1.98 J/cm²</td>
<td>3×</td>
<td>4.32–5.94 J/cm²</td>
<td>25 cm</td>
<td>2 × 15 W</td>
<td>&gt;4 log reduction</td>
<td>Decontamination methods did not significantly degrade filter performance at 300 nm particle size</td>
<td></td>
</tr>
<tr>
<td>Lindsley38</td>
<td>2015</td>
<td>n/a</td>
<td>120–950 J/cm²</td>
<td>1×</td>
<td>120–950 J/cm²</td>
<td>6.2 cm</td>
<td>2 × 15 W</td>
<td>27°C, 25% RH</td>
<td>No significance change in flow resistance, decreased penetration strength and strap breaking strength at very high exposure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mills11</td>
<td>2018</td>
<td>60–70 sec</td>
<td>1.1 J/cm²</td>
<td>1×</td>
<td>1.1 J/cm²</td>
<td>1 m</td>
<td>8 × 0.39 W/cm²</td>
<td>&gt;3 mean log reduction, straps less effectively decontaminated</td>
<td>21°C, 50% RH</td>
<td>Masks oversaturated with virus in physiologically relevant solvents sebum and mucin, ultraviolet germicidal irradiation levels fully decontaminated masks, statistically significant reduction in virus viability</td>
<td></td>
</tr>
<tr>
<td>Heimbuch12</td>
<td>2019</td>
<td>16–18 mW/cm²</td>
<td>60–70 sec</td>
<td>1.0–1.2 J/cm²</td>
<td>20×</td>
<td>20 J/cm²</td>
<td>1 m</td>
<td>8 × 0.39 W/cm²</td>
<td>&gt;3 mean log reduction, straps less effectively decontaminated, tested on SARS-CoV-1</td>
<td>22.5°C</td>
<td>No meaningfully significant effect on fit, air flow resistance, or particle penetration for 15 models tested, some models had significant effects due to donning/ doffing cycles</td>
</tr>
<tr>
<td>Card42</td>
<td>2020</td>
<td>0.1 mW/cm²</td>
<td>20 min/ side</td>
<td>240 mJ/cm²</td>
<td>60.6 cm</td>
<td>30W</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lowe9</td>
<td>2020</td>
<td>0.2 mW/cm²</td>
<td>5–6 min</td>
<td>300 mJ/cm²</td>
<td>10 ft</td>
<td>16x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: J, Joules; mJ, milliJoules; mW, milliwatts; RH, relative humidity; W, watts.
capsid. UV-C wavelengths (100–280 nm) have the highest decontamination efficiency because the maximum absorption wavelength is 260 nm and 280 nm for nucleic acids and proteins, respectively. Although little is known about the novel SARS-CoV-2, comparisons between SARS-CoV-2 and SARS-CoV-1 can be helpful in estimating conditions in which SARS-CoV-2 may persist. As mentioned, ssRNA viruses, like SARS-CoV-2, are the most susceptible type of virus to UVGI, which is important for understanding the potential success of UVGI treatment. SARS-CoV-2 can persist on plastic and stainless steel for up to 3 days, although at 72 hours virus titer had decreased 3-fold, and no viable SARS-CoV-2 was measured on cardboard after 24 hours. How long this virus can remain viable on PPE has yet to be studied and is likely a function of room humidity, contaminating fluids, mask materials, and construction. Because FFRs from different countries of origin use the same polypropylene material to create the mask filter, it is expected that UVGI would perform similarly on the various FFRs. As the following section details, the correct wavelength range is important for the treatment system and must be validated using a UV-C meter to ensure effective dosing in each setup.

**PROPOSED SOLUTION: INEXPENSIVE, SCALABLE, AND ACCESSIBLE UVGI SYSTEM FOR FFR DECONTAMINATION**

We propose a collaboration between public and private research institutions and health care facilities to increase the access to UVGI decontamination. Currently, UV-C bulbs required for UVGI are in limited supply; however, most laboratory biosafety cabinets (BSCs) are equipped with UV-C bulbs, and thousands of these bulbs are currently sitting idle, as research has been mainly halted to slow the spread of the virus. Additionally, many research institutions recommend other chemical means of decontamination, thus additionally leaving UV-C bulbs idle. These bulbs also may be available in existing health care facility settings (e.g., tuberculosis wards) where they can be used for decontamination via passive fixtures or air handling units. Although others have proposed decontaminating FFRs within these BSCs directly, this requires the masks to be transported to and from the research institutions and staff to run the decontamination cycles. Our approach advances previous work and that from Nebraska Medicine and allows medical sites to create their own decontamination system, using the idle bulbs, by retrofitting or creating custom light fixtures with off-the-shelf parts available at any retail hardware store.

**Building a Mask Decontamination System**

We have developed step-by-step instructions to create UVGI light fixtures that can be downloaded from our website: [https://www.gleghornlab.com/uvgi-decontamination](https://www.gleghornlab.com/uvgi-decontamination). There are multiple options for fixture assembly, depending on availability of materials, that include:

1. Modifying a premade commercial light fixture to fit new or existing UV-C bulbs:
   - Using an existing ceiling light fixture
   - Modifying a hanging ceiling fixture

2. Creating a custom fixture from off-the-shelf parts to fit new or existing UV-C bulbs

Because the number and type of bulbs available will vary, we created an easy to follow protocol to create this lamp using almost any common BSC UV-C bulb. The design plans we developed are not limited to BSC UV-C bulbs, and we provide details to adapt them to any UV-C bulbs. The digits printed on a UV-C bulb provide the information necessary to the end user to adapt these bulbs to create these custom UVGI lamps. For example, a bulb labeled “G30T8”: G stands for germicidal, 30 is the wattage, and T8 represents the size of the bulbs and pin geometry. For bi-pin bulbs, T5 has a 5/8 inch diameter, T8 has a 1-inch diameter, and T12 has a 1.5-inch diameter. This information is important to match to the correct fixture/bulb holder.

**Determining UV-C Exposure Time Required for UVGI Decontamination of FFRs**

Similar studies analyzing the dosages required for proper UVGI decontamination and reuse of FFRs report an optimal UV-C dose of approximately 300 mJ/cm². To determine UV dosage, design geometries, scalability, and N95 mask decontamination throughput of our designs, we compared a model of irradiation (measure of UV-C intensity per area) and compared it to measured values of UV-C irradiation from our modified fixture. UV-C intensity on a fixed plane from cylindrical source is nonlinear and is a function of bulb characteristics (length, wattage, radius) and distance to the bulb (Figure 1A). For a single bulb, the highest...
irradiation will be achieved along the midline and will rapidly decay with increasing irradiation width (Figure 1B). We used a UV meter (attenuation $\lambda = 254$ nm) to measure the irradiation area at discrete points using our modified fixture outfitted with a single UV-C bulb (G30T8 with 13.4 W UV output) taken from an existing BSC. The bulb was 88 cm long and placed 15 cm above the UV-C meter sensor. We tested areas along the midline and 20 cm away from the bulb’s midline in either direction.

**FIGURE 1.** (A) Schematic demonstrating the geometric variables that determine ultraviolet germicidal irradiation area. (B) Experimental measurements of ultraviolet intensities are well described by a theoretical model to calculate ultraviolet irradiation. (C) Ultraviolet-C intensities were measured in an irradiation area 88 cm x 40 cm and the midline was 15 cm from the bulb source. As expected, ultraviolet intensities decay with increasing distance from the bulb, with the lowest intensities measured at the corners of the array. (D) Cartoon of an example table top ultraviolet germicidal irradiation setup.

Abbreviation: UVGI, ultraviolet germicidal irradiation.
direction, with the sensor flat on the table surface at all locations tested. The highest irradiation was along the midline directly under the bulb and the UV intensities decay, as expected, with increasing irradiation width (Figure 1B, C). Importantly, these measurements demonstrate 2 points. Firstly, the model is a conservative prediction of the intensity distribution, and thus can be used to inform bulb configuration geometries described herein. Secondly, that the bulb intensity is not uniform along its length. Both of these details reinforce the need for user-developed designs and configurations to be validated with UV-C meter measurements to determine the minimum values of UV intensity over the exposure area to calculate exposure times needed in individual configurations.

We measured a new, unused N95 mask (model 7048, 3M) to be 13 cm wide, and therefore we could decontaminate an array of 6 x 3 masks (Figure 1D). We recommend using a reflective surface such as a stainless steel table or table coated with simple aluminum foil when irradiating on a table surface to maximize irradiation intensities due to reflection.

To calculate the UV-C exposure time, $t$, in seconds, $s$, needed for treatment, as:

$$ t = \frac{D}{I} \times 1000 $$

where $D$ is the desired dosage value and $I$ is the lowest measured irradiances value from the UV-C meter within the irradiation area.

For an array of 6 x 3 masks, using a single bulb, we use the lowest measured UV-C value over the exposure area (258 $\mu$W/cm², from Figure 1C) to determine the irradiation time. For our measured setup, the desired dosage ($D$) is 300 mJ/cm² for SARS-CoV-2, irradiance value ($I$) is 258 $\mu$W/cm², and exposure treatment time ($t$) is 1162.8 seconds or 19.4 minutes for each side of the mask.

$$ t = \frac{300}{258} \times 1000 = 1162.8 $$

Because the different UV-C bulbs able to be used in this proposed system will have parameters depending on the specifications such as size and age of the bulb, the end user of their own specific UVGI system will need to validate the UV-C values. This should be done using a UV-C meter to determine actual irradiation areas and irradiation times to properly decontaminate the FFRs.

**Optimizing UVGI Decontamination**

**Minimize Irradiation Deficits on Mask Surface**

It is important to note that masks placed too close together can create shadows that prevent effective UV decontamination. Additionally, when irradiating masks on a table surface without reflective surfaces, the dome shape of the masks inherently makes it difficult to decontaminate the surfaces on the far side of the mask relative to the position of the light source (Figure 2A, B). To combat this issue, we devised 2 solutions: (1) place a reflective “backboard” at the edge of the array, and (2) place a reflective wedge underneath the mask on the edge of the array to angle the masks toward the light source. A simple wall made of aluminum foil covered cardboard was placed at the edge of the irradiation area (Figure 2C, D). The UV light was able to be reflected off this surface to achieve greater irradiation on the far surfaces of the mask. Using this method, UV-C intensity measured on the mask surface was 63% of the intensity measured on the front mask surface. In the second approach, a reflective wedge was created by wrapping aluminum foil around a folded piece of cardboard (Figure 2E, F). This sloped surface decreases the effect of the domed shape of the mask by orienting the mask in the direction of the light source to decrease distance from the bulb and prevent shadows caused by the mask surface relative to the bulb position. Using this method, edge region intensities were 79% that of the measured intensities on the front surface of the mask. However, this modification, in general, places the mask closer to the light source, increasing the overall irradiation intensity by 298%. Such solutions are simple and important to reduce shadows from occurring. Calculated exposure times should be based on the lowest irradiation value at the edge of the irradiation area as measured on the edge of the mask.

**Reduce Distance to Fixture**

If we reduce the distance to the bulb, the measured irradiation on the masks will increase non-linearly and will reduce the time needed to achieve a dose of 300 mJ/cm² per side of the mask. Further increases in mask decontamination throughput could be achieved by adding a second light fixture which will allow exposure to both sides of the mask simultaneously, similar to the Nebraska Medicine configuration, reducing the total time for irradiation in half (Figure 2G, H).
FIGURE 2. Multiple configurations can improve irradiation on the mask surface. (A) Image showing the table top setup of our custom ultraviolet germicidal irradiation. White circles on the table top demonstrate the footprint for placement of masks to be decontaminated within an 88 cm x 40 cm irradiation area. A shadow is visible on the far surface of the mask (white arrow) indicating decreased exposure due to the curvature of the mask and angle of the ultraviolet rays. (B) The far surface of the mask on the outer rows of our irradiation area, the mask region furthest from the bulb, receives 15% of the ultraviolet intensity compared to the top of the mask. (C, D) By adding a reflective backboard of aluminum foil wrapped cardboard to the edge of the irradiation area, the ultraviolet rays can be reflected to increase the irradiation exposure to the far surface of the mask. (E, F) Alternatively, by creating a sloped surface for the masks at the edge of the array, mask surfaces can be more aligned toward the light source resulting in more uniform irradiation intensities across the surface. (G, H) Another decontamination setup includes suspending or hanging the masks. If there is access to a second, optional fixture, this allows for irradiation of both sides of the mask simultaneously.
**Add Additional Fixtures**

Additional light fixtures can also be placed in parallel to increase the total irradiation area. If placed in close enough proximity, the 2 light fixtures will create an irradiation overlap region which will increase the irradiation intensity to reduce the time needed to decontaminate (Figure 3A). By modeling the theoretical irradiation curves following thermal radiation view factors, we can determine the effects of adjusting the bulb spacing on the mask decontamination throughput capacity. The 100% irradiation value used in our model was delegated as the highest measured UV-C intensity that occurred in the center of the array (Figure 1C). For our single bulb setup, using an irradiation width of 20 cm, processing 18 masks in an array, the decontamination time is 19.4 minutes per side, and thus can decontaminate 27.9 masks/hour (Figure 3B). This time is determined by the lowest intensity measured in our array where the intensity is approximately 26% that at the midline. However, if we use a modified commercial light fixture outfitted with 2 UV bulbs, the bulbs are closely spaced resulting in high overlap in UV light and the peak intensity of UV is notably increased (Figure 3C, D). This geometry can either produce a larger irradiation width for the same exposure time from 20 cm to 29 cm to increase mask throughput (Figure 3C), or we can keep the irradiation width the same and yield a higher UV intensity (Figure 3D). A higher UV intensity decreases the exposure time required and increases mask throughput. Although counterintuitive, Figure 3D demonstrates highest mask throughput (59.6 masks/hour compared to 37.2 masks/hour) by keeping the irradiation field constant (i.e., fewer masks at once) and having a higher UV intensity to decrease exposure time. These measurements and calculations are for a single modified, existing, commercial 2-bulb light fixture.

Although having 2 bulbs in a commercial fixture increases throughput compared to 1 bulb, taking advantage of off-the-shelf components to build a custom fixture has significant advantages. Specifically, users can optimize bulb spacing to increase the uniformity of the irradiation field. If 2 bulbs are spaced at the 50% irradiation width intensity for 1 bulb, the UV intensities are additive, which creates a more constant irradiation field (Figure 3E, F). Using this configuration, the irradiation width is considerably increased from 20 cm to 34.5 cm for a mask exposure time of 19.4 minutes/side. This generates a mask decontamination throughput of 46.4 masks/hour (Figure 3E). However, similar to findings from the off-the-shelf fixture, keeping the irradiation area constant to take advantage of the higher UV intensities will increase mask throughput more substantially. This approach reduces decontamination time to 5.9 minutes/side and yields a throughput of 91.3 masks/hour for a system with lights only on 1 side of the masks (Figure 3F). This represents a greater than 3-fold increase in mask processing from 1 UVGI bulb. If a system is created with simultaneous exposure to front and back sides of the masks, using 2 UVGI bulbs on each side, a user could process 182.6 masks/hour. Choosing an optimal setup will all depend on the resources at hand, such as fixtures, bulbs, and treatment environments. The end user will need to validate the UV-C values, using a proper UV-C meter, once the desired configurations of UVGI bulbs are in place to determine actual irradiation areas and irradiation times to properly decontaminate the FFRs. A significant advantage of this system is the potential for parallelization of numerous UVGI light fixtures to simply scale the N95 mask treatment throughput.

**Approximate UVGI Fixture Assembly Costs**

This system leverages affordable components and resources already available and distributed across the country in a variety of different research centers. Using our approach to modify commercial fixtures from the hardware store, 1-bulb, 2-bulb, and 4-bulb UVGI light fixtures can each be constructed for less than US$25, US$30, and US$45, respectively. Each custom built fixture (Figure 4A, B) can be constructed for less than US$21 for a 2-bulb configuration and less than US$36 for a 4-bulb configuration. These costs do not include the cost of the bulb itself which we believe could be obtained by leveraging collaborations with research institutions to use UV-C bulbs that might otherwise be sitting idle. To implement this UVGI light source, a UV-C meter should be used to provide an accurate measurement of irradiance (μW/cm²) at the position of the FFRs will be placed away from the UVGI system. We recommend using a similar workflow and arrangement that Nebraska Medicine developed with their UVGI light towers. Both modified light fixtures or custom fixtures can be easily propped over the decontamination surface by simply resting on boxes (Figure 2A) or affixed to common items in a medical facility such as under a table or an intravenous pole to allow customization and adaptability of the.
FIGURE 3. Irradiation time can be optimized by placing 2 fixtures in close proximity. (A) The ultraviolet light from both fixtures will overlap to create an irradiation overlap region where the ultraviolet intensities are additive. (B) One bulb will create a single maximum irradiation peak; varying the distance between 2 bulbs creates different irradiation and decontamination scenarios which can be modeled to find the optimal conditions. (C, D) Two bulbs modeled in very close proximity; (E, F) modeled at an optimal distance can increase the irradiation area if the irradiation time is fixed or decrease the irradiation time by increasing the overall exposure within the original irradiation area.
FIGURE 4. An easy, affordable, custom ultraviolet germicidal irradiation system can be built with off the shelf parts. (A) Cartoon representation of our custom built ultraviolet germicidal irradiation fixture. (B) Finished product of our custom ultraviolet germicidal irradiation fixture. (C) Fixtures can be placed over tables or placed vertically by attaching it to health care equipment such as an intravenous pole. (D) Example arrays showing how a user might assemble a horizontal mask decontamination surface or a vertical surface which would allow for irradiation on both sides of the mask pending the availability of additional ultraviolet germicidal irradiation bulbs.
Inexpensive and Scalable Ultraviolet System for Filtering Facepiece Respirator Decontamination

Global Health: Science and Practice 2020 | Volume 8 | Number 3

Inexpensive and Scalable Ultraviolet System for Filtering Facepiece Respirator Decontamination

www.ghspjournal.org

594

on the measured UV-C output (\(\lambda=254\) nm) at a defined distance. When working with our proposed UVGI system, it is also important to use proper UV safety procedures, including avoiding any direct eye or skin exposure to the UV light source.

Whereas these instructions are illustrated using materials available in the United States, this concept has global application. The same bulbs are used in research BSCs, medical settings, and increasing water treatment and food safety applications throughout the world. The ability for HCWs to quickly decontaminate PPE on site during any health care emergency could be an important measure to reduce the spread of infectious diseases, especially when time and other resources are limited.

Funding: Although this specific project was not directly funded by specific grants, we would like to thank the funding agencies of our other projects including the National Institutes of Health, the National Science Foundation, and the Bernard Canavan Award.

Competing interests: None declared.

REFERENCES


© 2020 Global Health: Science and Practice
30. Halperin E. Battelle cleared to sterilize N95 masks at max capacity, doi.org/10.9745/GHSP-D-20-00218

28. Centers for Disease Control and Prevention. Implementing filtering facepiece respirator decontamination


Using Patient-Reported Outcome Measures to Promote Patient-Centered Practice: Building Capacity Among Pediatric Physiotherapists in Rwanda

Monika Mann,a Ines Musabyemariya,b Linn Harding,c Ben Braxleyd

Key Findings

After completing the training:

- Ninety-two percent of respondents stated that they felt either “quite” or “very” confident in their abilities to establish functional goals.
- Seventy percent of respondents stated that they continue to use patient-reported outcome measures either “a lot” or “quite a bit” 26 months after the close of the grant.

Key Implications

Program managers should consider:

- Targeting stakeholders from a range of health system levels to facilitate systemic and institutional support of practice changes as well as maximize adoption, penetration, and sustainability of new skills and concepts
- Adopting culturally appropriate outcome measure tools
- Using multimodal strategies for capacity building

ABSTRACT

Background: Patient-reported outcome measures (PROMs) assess disability and progress toward functional goals while promoting patient-centered practice. They can be used by health professionals in any specialty and in a multitude of settings. This study reviews implementation strategies and lessons learned in a capacity-building program that took place with pediatric physiotherapists in Rwanda.

Methods: Use of PROMs and patient-centered practice were integrated into 4 consecutive continuing professional development courses offered to 164 participants in Rwanda. We sought to identify a simple generic measure with proven validity in cross-cultural settings. The Patient-Specific Functional Scale was chosen due to its ease of use and ability to measure change in a wide range of patient conditions. Didactic classroom training and clinical site visits were 2 essential pedagogical elements of the capacity-building strategy. Site visits allowed for evaluation of skill levels and facilitation of knowledge transfer to patient care settings. Unique pairs of Rwandan colleagues were trained to serve as coteachers in each course to maximize sustainability of new techniques. This study presents data on a subset of 65 participants who completed a 48-hour pediatric rehabilitation course.

Results: After classroom instruction, 78% of participants were observed independently determining functional limitations with their patients. Additionally, pre- and post-tests indicated that therapists substantially increased their understanding of patient-centered practice after attending courses. Interviews conducted 26 months after the conclusion of the project revealed mixed success in sustainability of the use of PROMs, although perceived confidence remained high.

Conclusion: Challenges in long-term sustainability of new practices call attention to the need to target not only clinicians when introducing new methodologies, but also the Ministry of Health, hospital administration, and university faculty. Lessons learned from this study may be useful to other medical professionals planning capacity-building programs in low- and middle-income countries.

INTRODUCTION

Tracking health care outcomes is integral to assessing effectiveness and efficiency within health systems. Historically, the data have been based on medical tests of body function and structure performed by a clinician. Patient-reported outcome measures (PROMs) were developed to focus on activity limitations and participation restrictions reported by the patient (Figure 1).
Treatment planning and assessment can thereby focus on the outcome goals of the patient rather than solely the clinician’s objectives. PROMs not only take into account the patient’s experience, but also tend to be oriented toward quality-of-life measures including function and ability to participate in society.1

Standardized and validated instruments to measure outcomes are utilized to both guide treatment and determine its functional impact.3 It is crucial to integrate PROMs as components of the data routinely collected and analyzed to promote patient-centered practice.4 Although PROMs can be used by health professionals in any specialty and in a multitude of settings, this report reviews implementation strategies and lessons learned during a capacity-building program to increase the use of PROMs and promote patient-centered practice among pediatric physiotherapists in a low-resource country.

PROGRAM DESCRIPTION
The Advancement of Rwandan Rehabilitation Services Project (ARRSP) was a 27-month program funded by the United States Agency for International Development and implemented by Health Volunteers Overseas from March 2013 until May 2015. Key goals of the program were to upgrade rehabilitation standards in Rwanda and improve the quality of services provided. This aim was in alignment with Rwanda’s Third Health Sector Strategic Plan, July 2012–June 2018, in which one of the strategies listed was to “train health workers on control, prevention, and treatment of injuries and disabilities.”5

The ARRSP offered successive continuing professional development courses to practicing physiotherapists. Content specialists were recruited from the United States to teach a series of courses on topics that had been selected by a steering committee made up of Rwandan rehabilitation professionals. The committee also chose 2 different Rwandan physiotherapists to be co-instructors for each course. Preference was given to physiotherapy faculty members so that they could integrate new concepts and practices into their teaching of students at the university. During the grant period, there were 4 physiotherapy faculty members in the country, and all participated in the courses. Rwandan co-instructors received intensive training on PROMs, course content, and patient-centered practice. This study focuses on the participants of the pediatric rehabilitation course since a follow-up assessment was conducted on this subgroup 26 months after the conclusion of the grant to determine sustainability of introduced skills and concepts.

Courses ranged from 36 to 48 hours in length and were offered to 2 or 3 cohorts at a frequency of one weekend class session per month per cohort. By the last class session each month, Rwandan co-instructors were responsible for the majority of teaching. At the time the courses were given, records from the Rwanda Allied Health Professionals Council indicated that there were 142 registered physiotherapists employed in direct patient care in the country. One hundred sixty-four therapists attended at least one of the courses included in this study, with most participants attending 2 or 3 continuing professional development offerings (Table 1). This indicates that virtually all practicing physiotherapists as well as some therapists who did not work directly in patient care attended at least one of the multi-
weekend offerings. About 22 physiotherapists worked with a population that was at least 75% pediatric. Most other therapists were generalists and saw a mix of adult and pediatric patients.

All courses incorporated didactic classroom teaching as well as training in clinical settings. An essential element in both environments was the use of PROMs and patient-centered clinical decision making to facilitate patients’ return to optimal function. To achieve this, we set out to select a contextually appropriate, generic PROM that could be used across all courses.

**Outcome Measure Selection**

Since the use of outcome measures has been identified as an essential component of best practices in physical therapy,6,7 an objective of the ARRSP was to instruct Rwandan physiotherapists in the use of an appropriate PROM that would promote patient-centered practice. The use of outcome measures was a new concept among the majority of physiotherapists in Rwanda, so we sought to identify a simple, generic measure that could be used across multiple diagnoses. Both the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) and Patient-Specific Functional Scale (PSFS) have been shown to be valid in cross-cultural settings and can be used to measure change in a wide range of patient conditions.8–13 Therefore, these were the 2 PROMs considered for use.

Studies have demonstrated that consideration of local, contextual factors is often essential to the long-term success of a project.14,15 This is especially salient when working outside of one’s usual environment. During our early meetings with University of Rwanda faculty and administration, we were advised that the Rwandan culture tended toward oral communication and that processes involving extensive reading or writing could represent barriers to the acceptability, adoption, and sustainability16 of a PROM. Additionally, in many clinics and hospitals, the resources necessary to produce a lengthy PROM, such as the WHODAS 2.0, were not reliably available. This barrier cast doubt on the appropriateness and feasibility of the WHODAS 2.0 for use in everyday clinical practice. The PSFS can be administered verbally with little reading required, and minimal space is needed to record scores (Figure 2).

An additional strength of the PSFS is its inherent patient-centeredness. Rather than using predefined activities that might lack a local context, the PSFS calls on patients to identify and assign a rating to activities that they have difficulty completing as a result of their impairment. In this manner, the scale is tailored toward activities that are relevant to each patient and a separate pediatric version is not necessary. For infants, toddlers, and nonverbal children, functional deficits and goals can be determined by a proxy, such as parents. The measure has demonstrated reliability and validity for use in impairments of the lower extremity, upper extremity, and the cervical and lumbar spine.17 Additionally, the use of PROMs in a pediatric population has been described by a number of authors.18–21 Taking these attributes into consideration, we selected the PSFS for use throughout the ARRSP.

### Capacity-Building Strategies

The ARRSP sought to build capacity through training clinicians and rehabilitation faculty in PROMs and patient-centered practice. The ARRSP sought to build capacity through training clinicians and rehabilitation faculty in PROMs and patient-centered practice. Once we determined that the PSFS would be the PROM of choice, the course instructors and project coordinator began an iterative process of refining implementation techniques for capacity building in both use of the measure and improvement of patient-centered practice. Because 4 unique instructor teams each taught successive courses covering a different topic in rehabilitation, capacity building evolved as the unique needs of each cohort became apparent. A 3-pronged approach

<table>
<thead>
<tr>
<th>Course Name</th>
<th>Cohort Groups</th>
<th>Classroom Hours per Cohort</th>
<th>Class Sessions per Cohort</th>
<th>Number of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Therapeutic exercise</td>
<td>3</td>
<td>48</td>
<td>4</td>
<td>81</td>
</tr>
<tr>
<td>Neurological rehabilitation</td>
<td>3</td>
<td>48</td>
<td>4</td>
<td>90</td>
</tr>
<tr>
<td>Pediatric rehabilitation</td>
<td>2</td>
<td>48</td>
<td>4</td>
<td>65</td>
</tr>
<tr>
<td>Leadership institute</td>
<td>2</td>
<td>36</td>
<td>3</td>
<td>69</td>
</tr>
</tbody>
</table>
was used to maximize acceptance, adoption, penetration, feasibility, and sustainability of the measure.

Three-Pronged Approach to Capacity Building

1. Modeling and Reinforcing PSFS Technique and Patient-Centered Practice

Course instructors used the World Health Organization’s International Classification of Functioning, Disability and Health (ICF)\(^2\) as a model for teaching and reinforcing the concept of functioning and functional goals. Classroom time was scheduled to allow for multiple repetitions of modeling appropriate PSFS utilization. Didactic teaching included lecture, case study discussions, role-playing, and hands-on skills training. Emphasis was placed on using results of the PROM to design patient-centered treatment plans and objectively tracking progress toward functional goals. Classroom work in small groups allowed for discussion on how to revise treatment plans based on changes in PSFS scores during reassessments. Instructors often referenced cases observed during clinic visits to provide culturally appropriate examples of PSFS use and relevance to patient-centered practice.

2. Distributing a Standardized Booklet for Ease of Use by First-Generation PROM Adopters

In an attempt to maximize adoption and feasibility of the PSFS, a booklet was drafted and distributed to participants that included instructions and rationale for its use. The booklet also contained a completed sample PSFS and 3 blank forms. PSFS data collected via site visits and booklets were presented in the classroom to facilitate discussions on how results can be used to assess and progress treatment.

3. Evaluating the Use of the PSFS and Patient-Centered Practice During Clinical Visits

The inclusion of clinical site visits was an integral component of the ARRSP and was thought to be
essential for achieving maximal adoption and penetration of the PSFS and patient-centered practice. Site visits provided an opportunity to offer direct clinical mentoring and observation. This allowed instructors to evaluate skill levels and reinforce classroom lessons to help facilitate knowledge transfer to patient care settings. A site visit checklist was utilized to evaluate clinical reasoning, documentation of functional problems, patient-centered practices, and implementation of the PSFS. Formative feedback was offered to those who needed assistance completing the measure or had difficulty relating it to patient-centered practice. Course participants typically invited colleagues to engage in clinical visits so that the number of rehabilitation professionals influenced by the visits substantially exceeded the number formally attending classes.

### RESULTS

We used 6 measures to assess the success of the project:

1. Pre- and post-tests for each course
2. Clinical observations
3. Participants’ written assessments of each course
4. End-of-grant written assessments
5. End-of-grant interviews with course participants
6. Interviews and survey conducted 26 months after the conclusion of the ARRSP

### Pre- and Post-tests

Pre- and post-tests were a part of every course. Item analysis of the pre- and post-tests indicated that participants’ understanding of PROMs and patient-centered practice increased after attending courses. An example of a question included in the pre- and post-test of the pediatric rehabilitation course is shown (Box). Supplement 1 contains the full pre- and post-test questionnaire. On the pretest, 55% of the participants (N=66) selected option B, the correct answer. On the post-test, 94% (N=61) chose the correct answer, demonstrating an increased understanding of the value of the PSFS.

### Clinical Observations

Course instructors evaluated the participants’ ability to identify functional deficits during structured clinical observations. This skill was new and was not previously a routine part of clinical practice. In the pediatric rehabilitation course, observations were completed on 60 participants. As can be seen in Table 2, 78% independently determined functional problems during clinical observations, indicating successful transfer of didactic knowledge to clinical practice.

### Course Assessments

Before this project, therapists did not routinely establish or progress functional goals. Instead the typical practice was to write general goals based on symptoms such as “reduce pain” or “improve strength.” As part of the pediatric course assessment (Supplement 2), participants were asked to rate their perceived competency in various indicators related to functioning (Table 3). As can be seen, 92% of respondents stated that they felt either “quite” or “very” confident in their abilities to establish functional goals and 97% stated that they were either “quite” or “very” confident in “progressing functional, meaningful treatment activities.”

### End-of-Grant Assessments

Assessments were distributed at the closing ceremonies of the project (Supplement 3). Fifty-five participants who attended at least one of the ARRSP courses completed the assessments. This represents 35% of the total number of unique...
Table 3. Perceived Competency in Various Indicators Related to Functioning at the Conclusion of the Pediatric Rehabilitation Course

<table>
<thead>
<tr>
<th>Thinking about the last 5 patients you saw in your workplace last week, how confident were you in:</th>
<th>Very Much</th>
<th>Quite a Bit</th>
<th>Somewhat</th>
<th>A Little Bit</th>
<th>Not at All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifying activity limitations (N=63)</td>
<td>53 (84%)</td>
<td>9 (14%)</td>
<td>1 (2%)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Establishing functional goals (N=62)</td>
<td>28 (45%)</td>
<td>29 (47%)</td>
<td>5 (8%)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Selecting functional, meaningful treatment activities (N=55)</td>
<td>28 (51%)</td>
<td>26 (47%)</td>
<td>1 (2%)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Progressing functional, meaningful treatment activities (N=66)</td>
<td>29 (44%)</td>
<td>35 (53%)</td>
<td>1 (2%)</td>
<td>1 (2%)</td>
<td>0</td>
</tr>
</tbody>
</table>

*Not all participants answered every question resulting in a variation in the number of respondents.

Table 4. Self-Rated Improvement in Patient-Centered Practice Measured at the End of the Grant Period

<table>
<thead>
<tr>
<th>As a result of attending the Advancement of the Rwandan Rehabilitation Services course, how much do you feel your evaluation and treatment of patients has improved for the following: (N=55) (Likert 0–10 Scale used)</th>
<th>Responses</th>
<th>Mode</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using outcome measurements</td>
<td>7</td>
<td>7.8 (1.4)</td>
<td></td>
</tr>
<tr>
<td>Adjusting treatment based on patient improvement</td>
<td>9</td>
<td>8.2 (1.1)</td>
<td></td>
</tr>
<tr>
<td>Evaluating functional activities</td>
<td>8</td>
<td>8.4 (1.1)</td>
<td></td>
</tr>
<tr>
<td>Setting functional improvement goals</td>
<td>8</td>
<td>8.3 (1.2)</td>
<td></td>
</tr>
<tr>
<td>Clinical decision making</td>
<td>8</td>
<td>7.9 (1.1)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: SD, standard deviation.

Course participants. As part of the assessment, they were asked to rate their perceived improvement in various aspects of patient-centered practice and clinical reasoning (Table 4). A Likert scale was utilized to measure improvement as follows: 0 = not at all, 5 = somewhat, and 10 = a great deal. As can be seen in the table, the mode of perceived improvement in all indicators was 7 or greater on the Likert scale.

End-of-Grant Interviews

Near the end of the grant period, in-depth interviews were held with physiotherapy department managers from 4 major hospitals who had participated in ARRSP courses. Because it was not feasible to interview every course participant, we chose managers since they could offer insight to practice changes of therapists who worked in their departments. A written semistructured interview guide was used and responses were simultaneously recorded in writing. All responses to questions about outcome measures were positive. Below is a typical sampling of responses:

*We never used functional outcome scales previously. We now use the patient-specific functional scale. It helps us in formulating goals: Now we focus more on needs of the patient rather than just the physiotherapist’s expectations. Using the PSFS has also helped us with our clinical decision making. If a patient is progressing, we continue with the same treatment, but if they are not, then we change the treatment.* —Respondent

*The staff now uses the PSFS to help assess patients and form goals. We now measure goals more quantitatively, like timing how long a patient can stand.* —Respondent

*The physios now use the PSFS with each initial evaluation. It helps with discharge planning and progression of the patient.* —Respondent

Interviews and Survey Conducted 26 Months After the Conclusion of the ARRSP

In the summer of 2017, a postgrant evaluation was conducted 26 months after the conclusion of the last course offering. An online survey was sent to the 65 participants of the pediatric rehabilitation course (Supplement 4). Forty-three people (66%) completed the survey. As can be seen in Table 5, 70% of respondents stated that they now used the PSFS either “a lot” or “quite a bit.”
Additionally, the survey asked participants to rate their perceived competency in the same indicators related to functioning that they self-rated in the course assessment that was filled out on the last day of the course. Table 6 compares responses given at these 2 points in time. There was no more than a 10% difference in perceived competency at the 2 points in time, suggesting good sustainability of skills and concepts learned.

In 2016, the Rwanda Ministry of Health outlined new guidelines for data collection stating that it is essential to collect accurate data to track patient outcomes. Performance standards are required in all 42 district hospitals in the country to attain re-accreditation. Monitoring clinical outcomes is designated as critical to improving quality and safety of care.

Although these changes are positive, there are still obstacles to the long-term adoption of PROMs. Interviews and clinical observations held 26 months after the conclusion of the ARRSP revealed that PROMs were not regularly completed. Two frequently cited reasons included lack of promotion of the measures by department managers and the feeling among therapists that they were too time-consuming.

Two further obstacles were noted at the institutional level. Although we supplied participants with a number of PSFS booklets for use in the workplace, there were barriers to producing more. Copiers were not available in physiotherapy departments and blank sheets of paper were often in short supply. Additionally, initial evaluations were often stored centrally within hospital records and not easily accessible to therapists after the evaluation had been completed. This storage issue created an obstacle to intermittent monitoring to assess changes in function and modify treatment programs accordingly.

**DISCUSSION**

Results indicate that the use of the PSFS reached encouraging levels of acceptability, adoption, feasibility, and penetration at the end of the grant period. Additionally, participants demonstrated greater understanding and utilization of patient-centered practice techniques. The postgrant evaluation found that perceived confidence in establishing functional goals; selecting functional, meaningful treatment activities; and progressing those activities remained high 26 months after the grant period ended. However, the evaluation also revealed obstacles that limited optimal sustainability of the use of PROMs. Both institutional and clinical practice challenges were cited. The challenges in achieving long-term sustainability highlight difficulties that can be encountered when introducing new methodologies in low-resource settings. There must be a solid foundation upon which to introduce new practice standards. Accomplishing this is difficult without
support at every level of the health system. In this project, we used the capacity-building model illustrated in Figure 3 in which the targets for interventions were practicing clinicians and rehabilitation faculty.

In retrospect, we think the model in Figure 4 would have been better. By eliciting support from multiple levels of the health system including the Ministry of Health, hospital administration, and physiotherapy department managers, we believe that sustainability of new methodologies and practice techniques could have been enhanced.

**Limitations and Lessons Learned**

Although the Rwandan steering committee requested that ARRSP courses emphasize clinical decision making, the majority of physiotherapists in Rwanda were not familiar with outcome measures and patient-centered practice. Therefore, there were uneven perceptions of need for change and motivation to change. Some therapists understood that using PROMs was an essential aspect of clinical decision making and were motivated to use them. But others did not fully integrate them into their practices.

Because the end-of-grant assessment was distributed at the closing ceremonies, it was only available to those in attendance, which was approximately 35% of course participants. If we could do it over, we would send out an electronic version of the assessment to reach more participants.

The postgrant assessment was performed 26 months after the grant closing and only included participants who attended the pediatric rehabilitation course. This was due to the interest,
availability, and motivation of one of the Rwandan co-instructors in sampling this subset. In hindsight, we would have implemented the postgrant assessment 12 months after the final course and included participants from all ARRSSP courses to assess the levels of success across all participants. Additionally, the pediatric group was challenging since many of the patients had long-term congenital conditions that do not respond quickly or completely to physical therapy interventions.

CONCLUSION AND RECOMMENDATIONS

We recommend that groups planning similar capacity-building endeavors consider the model presented in Figure 4 to target stakeholders from a broad range of health system levels. We believe that this approach will better facilitate systemic and institutional support to maximize adoption, penetration, and sustainability of new skills and concepts. We hope that methodologies utilized and lessons learned from this study are useful to other medical professionals planning capacity training in low-resource settings.

Acknowledgments: The authors wish to thank Kathryn Clark, Cara N. Whalen Smith, Lori Kahls, Egide Kayonga, Jean Damasen Gasherekuba, Jeanne Kagwisa, David Tumusilime, Linda James, Nancy Kelly, and Meng Xiong.

Funding: The Advancement of Rwandan Rehabilitation Services Project was implemented by Health Volunteers Overseas. Funding was made possible by the generous support of the American people through the United States Agency for International Development (USAID). Course content is the responsibility of Health Volunteers Overseas and does not necessarily reflect the views of USAID or the United States Government.

Competing interests: None declared.

REFERENCES


www.ghspjournal.org
The most useful resource reported was Global Surgery 2030, along with other publications, data collection tools, books and training manuals, and a documentary.

This list could serve as a starting point for individuals interested in global surgery and be supplemented with resources advocating for global surgery from clinical, population health, or policy perspectives.

THE GROWING FIELD OF GLOBAL SURGERY

Global surgery is defined as:

> an area of study, research, practice, and advocacy that seeks to improve health outcomes and achieve health equity for all people who need surgical and anesthesia care, with a special emphasis on underserved populations and populations in crisis.

The need is great. Surgical disease is among the top 15 causes of disability, and surgical conditions account for up to 30% of total disability-adjusted life years (DALYs) lost worldwide—with the greatest need in low- and middle-income countries (LMICs). Surgery has been shown to be highly cost-effective when compared with standard global health interventions.

The transition from the Millennium Development Goals to the Sustainable Development Goals has ushered in a new era for the global surgery community. Sustainable Development Goal 3, to “ensure healthy lives and promote well-being at all ages,” emphasizes health system strengthening and universal health coverage. The provision of available, accessible, safe, timely, and affordable surgical and anesthesia care is identified as an integral component of a functional health system in countries at all levels of economic development and as essential to achieving universal health coverage. In addition, the importance of increasing education, safety, and capacity for the provision of surgical, anesthetic, and obstetric care is highlighted by several global health and development agencies and policy makers, including the World Bank and the World Health Organization (WHO).

As a result, the emerging field of global surgery has increased in priority among health practitioners, including nonphysician surgeons and anesthetists, researchers, and students. Evidence of this prioritization includes a shift toward incorporating surgical care as an integral part of global health systems strengthening in LMICs that has occurred and will likely continue to grow in importance within global health agendas. Lastly, interest in the field from an academic research standpoint is evidenced by the increase in peer-reviewed publications. Between 2005 and 2015, research publications in the field of global surgery increased from approximately 570 articles in 2005 to more than 4,000 articles published in 2015, according to PubMed.

Because of the growing interest in global surgery, momentum in this emerging field, and the importance of global surgery in the training of health professionals, we aimed to summarize the top resources in global surgery to orient readers to the field. We undertook a 2-stage process to identify and select the top 10 resources in global surgery.

In the first stage, we convened a team of global surgery leaders, including persons with decades of experience in global surgery and emerging leaders. The team included 6 global surgeons from a variety of surgical specialties and training and 2 public health professionals. Evidence of this prioritization includes a shift toward incorporating surgical care as an integral part of global health systems strengthening in LMICs that has occurred and will likely continue to grow in importance within global health agendas.

In the first stage, we convened a team of global surgery leaders, including persons with decades of experience in global surgery and emerging leaders. The team included 6 global surgeons from a variety of surgical specialties and training and 2 public health professionals. A diverse authorship team was deliberately selected so that the nominated resources would transcend disciplines, career stages, and setting. The authorship team also included representation from varying career stages, including 2 students, 2 residents, 3 practicing surgeons, and 2 academic researchers. Lastly, the team included
Twenty-one resources were nominated de novo, including 10 peer-reviewed articles, 4 books or monographs, 3 data collection tools, 2 development manuals, and 2 advocacy pieces, comprising a documentary film and a newspaper opinion-editorial article.

The second stage involved a collaborative crowdsourcing initiative using social media and email distribution lists over a 1-month period. Using the social media and email platforms, we aimed to collect information from persons involved in global surgery from several perspectives, including surgeons, anesthesiologists and anesthetists, obstetricians, and public health students. We created a short survey that contained the 21 resources from stage 1 and asked participants to select their top 10 resources from this list. Participants were also asked to nominate other resources related to global surgery through short-answer responses. After the 1-month period, ratings and open-ended responses were compiled and tabulated, resulting in the nominated top 10 resources in global surgery.

## GLOBAL SURGERY TOP 10 RESOURCES

The top 10 resources included a wide variety of items including 3 publications, 3 data collection tools, 3 books or manuals, and 1 documentary film (Figure 1).

### Publications

- Surgery and Global Health: A View from Beyond the OR (2008) is widely regarded as the impetus

### Datasets/Checklist

- World Health Organization Surgical Safety Checklist
- Global Burden of Disease data
- World Development Indicators

### Books/Manuals

- Disease Control Priorities, 3rd edition. Volume 1: Essential Surgery
- Global Surgery: An Introduction
- Surgical Care Systems Strengthening: Developing National Surgical, Obstetric and Anaesthesia Plan

### Video

- The Checklist Effect

---

**FIGURE 1.** Top 10 Resources in Global Surgery
for global surgery to become a major global health emphasis. Written by 2 key leaders in global health, Dr. Paul Farmer and Dr. Jim Kim outlined the need for surgery in LMICs and defined global surgery as the “neglected stepchild of global health.”

Global Surgery 2030, a report by The Lancet Commission on Global Surgery (2015), repeatedly topped lists as the most useful and impactful resource. Launched in 2014, the Commission comprises a multidisciplinary, international team of commissioners, collaborators, and advisors from more than 110 countries and 6 continents. Its report, released in the following year, proposed a plan of action for this scale-up over the next 15 years. The data created by the Commission’s researchers have been included in the World Bank Development Indicators data collection and served as a launching point for the development of national surgical plans and the collection of primary global surgery data.

In 2014, Grimes and colleagues wrote Cost-Effectiveness of Surgery in Low- and Middle-Income Countries: A Systematic Review to evaluate the cost-effectiveness of surgical interventions that could be available at district hospitals in LMICs. The authors argued that many surgical interventions are as cost-effective, life-saving, and disability preventing as traditional global public health measures (e.g., vaccination, bed nets for malaria prevention), thus debunking the myth that the provision of surgery is too expensive for low-resource settings.

Datasets/Data Collection Tools
As part of the Safe Surgery Saves Lives initiative formed by the World Alliance for Patient Safety and WHO, the WHO Surgical Safety Checklist was created to assist surgical teams in reducing avoidable surgical site infections and preventable deaths. The checklist, which has been implemented in over 4,100 hospitals and still actively used in more than 1,790 hospitals, is associated with reduced complications, infection, and mortality rates, improved team communication, and increased economic benefits.

The second data collection tool resource is the online Global Burden of Disease (GBD) data (2017) repository and visualization tools. The GBD Results Tool, hosted by the Institute for Health Metrics and Evaluation, allows for a comprehensive overview of disease morbidity and mortality around the world and comparison between countries, risk and causal factors, and patient demographics.

The final data collection tool resource includes the 6 core indicators reported by the Lancet Commission report that were incorporated in the World Bank Development Indicators in 2015. The incorporation of surgery-specific health indicators in the World Bank database, the largest collection of health, development, and economic data from member states, represents the first attempt to comprehensively gather primary surgical data on an international scale. In addition, the open-access nature of the data encourages accessibility to a wide range of users rather than only academics or those with sufficient economic resources.

Books and Manuals
The third edition of the Disease Control Priorities (DCP3) series provides “the most up-to-date evidence on cost-effective interventions to address the burden of disease in low-resource settings.” Volume 1 of DCP3 is dedicated to Essential Surgery and evaluates 44 emergency and essential surgical procedures.

Global Surgery: An Introduction, a booklet written by Dr. Dominique Vervoort in 2017, provides an introduction on the basic concepts, history, and future steps of global surgery and related topics, targeting an audience new to the field of global surgery.

Lastly, following the movement toward National Surgical, Obstetric, and Anesthesia Plans (NSOAPs) in some sub-Saharan African countries and the launch of Zambia’s NSOAP, the WHO produced the Surgical Care Systems Strengthening: Developing National Surgical, Obstetric, and Anaesthesia Plans to highlight the importance of NSOAPs and the progress made in selected countries around the world.

Documentary
Finally, The Checklist Effect, a 2015 documentary by the Lifebox Foundation directed by Lauren Anders Brown, follows surgical health professionals in Haiti, Uganda, Mongolia, Guatemala, and Moldova. The documentary has been screened and awarded at several international global health film festivals and translated into 10 languages by student members of the International Student Surgical Network, an international nonprofit organization, consisting of medical students, residents, and young doctors interested in global surgery.

ADDITIONAL RESOURCES
Due to the increased interest in global surgery among academic institutions and LMIC health
care professionals, there is not only the need for a collection of first-step resources in the field but also the need for thoughtful guidelines to improve surgical care in LMICs and low-resource settings through multidisciplinary approaches.

There are many additional resources, including publications, partnerships, training tools, and ongoing studies on global surgery that are moving the field forward in tremendous ways. This list is not meant to be comprehensive or final, but rather a starting point for students, professionals, clinicians, and others interested in global surgery. Initially, the focus of interest in global surgery has solely been on providing clinical care to persons with surgical conditions. However, the field has broadened to include public health, health economics, health services, and policy research perspectives, thus promoting the need for global surgery resources to encompass both the individual-level and population-level needs, while contextually incorporating locally-driven viewpoints. Building upon the top 10 resources, other key supplementary resources from a socioecological viewpoint are described below (Figure 2).

**Clinical Perspective**

On a clinical level, the field of global surgery is seeing momentum build around clinical guidelines for care in resource-constrained settings that are relevant to practitioners in LMICs. Some clinical practice resource examples include *Primary Surgery, Principles of Reconstructive Surgery in Africa, Paediatric Surgery: A Comprehensive Textbook for Africa*, SUGAR PEARLS (Procedural Education for Adaptation to Resource Limited Settings), and many more. In addition, locally driven surgical skills and training resources, such as *Global Surgery and Anesthesia Manual* and *Non-Technical Skills for Surgeons System Handbook* have also been made available to health practitioners in LMICs.

Additional effects toward trauma care improvement, including prehospital and emergency room care, have also been emphasized from a full health care system perspective. Although prehospital chains (e.g., through ambulance services or clear referral systems) and emergency services remain scant and under-addressed in global health, they form critical barriers in access to care by increasing the delay in reaching care. For example, emergency medical services systems are in place in only one-third of the countries in sub-Saharan Africa, only covering less than 10% of the population. Lastly, but importantly, nonphysician clinicians and nonspecialist physicians can also assist in closing the gap in needed surgical services at first-level hospitals. A recent review by Falk et al. emphasized the variation in recognition, training pathways, and procedural support that nonspecialist and nonclinician providers engage in through task-sharing models. Task sharing has accordingly become a predominant approach in surgical care delivery in many LMICs with formalized training pathways in, for example, Mozambique and Sierra Leone.

As a recommendation for the development of future resources, international collaborative efforts and partnerships (whether researchers in both HICs and LMICs or researchers in LMICs among themselves) can guide the creation of global surgery resources that cater to the needs of local stakeholders among practitioners in LMICs. Trainees and surgeons in LMICs must be fully involved in these endeavors and their need for more specific and targeted clinical resources addressed as well. Though already exposed to care in resource-constrained settings, trainees in LMICs need clinical

---

**FIGURE 2. Socioecological Model of Additional Global Surgery Resource Recommendations**

<table>
<thead>
<tr>
<th>POLICY</th>
<th>POPULATION</th>
<th>CLINICAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>• National Surgical, Obstetric, and Anesthesia Plans</td>
<td>• Longitudinal data collection studies</td>
<td>• Locally-driven clinical guidelines</td>
</tr>
<tr>
<td>• Resolutions</td>
<td>• Validated data collection tools</td>
<td>• Training manuals</td>
</tr>
</tbody>
</table>

Global Health: Science and Practice 2020 | Volume 8 | Number 3

609
guidelines such as those previously cited as they prepare to care for patients in their local low-resource settings. Reciprocity with collaborative partners must be ensured, sharing clinical experience from HICs and learning new approaches to surgical disease already practiced by local surgeons.35

### Population Health Perspective

On a population health level, collaborative efforts resulting in the collection of primary epidemiologic data are needed. Several large, multinational cohort studies on surgery, such as GlobalSurg, Global PaedSurg, COVIDSurg and other national and international databases on trauma care, neurosurgery, pediatric surgery, and other specialty areas, continue to collect primary data in LMICs, which has resulted in several profound publications highlighting the global surgery need from an epidemiologic standpoint.

In addition to the longitudinal cohort studies, there are several data collection tools available for population-level data collection, including the aforementioned tools used in the Global Burden of Disease tools and the World Bank Development Indicators. Other data collection tools include the Surgeons Overseas Assessment of Surgical Need, the Surgeons Overseas PIPES—Personnel, Infrastructure, Procedure, Equipment, and Supplies—Assessment tool, Surgeons Overseas Pediatric PIPES, the Global Assessment of Pediatric Surgery, and the World Health Organization Tool for Situational Analysis to Assess Emergency and Essential Surgical Care.

Recommendations for future resources from a population health perspective include additional longitudinal studies that are contextually applicable to specific populations. Validated questionnaires assessing a more granular approach to measure the burden of surgical conditions, facility-level tools to measure quality of care, and qualitative data collection tools aimed at overcoming barriers to surgical care are a few examples of further tools to be developed.

### Policy Perspective

Strong support for strengthening emergency and essential surgical and anesthesia care has been advocated as an essential component of universal health coverage and was recognized by the World Health Assembly (WHA) as resolution WHA68.15. In 2017, another resolution was passed (WHA70.22) that requires member nations to provide biennial reports on the strengthening of emergency and essential surgical care and anesthesia as a component of universal health coverage as requested in resolution WHA69.11. This series of global events at the policy level of health care highlight the role that policy makers and advocates of global surgery have on the international health agenda.

The promotion of recommendations to develop national surgical plans have propelled many countries to engage local stakeholders at various levels of health care and policy to begin formulating NSOs. Many organizations and institutions have been involved in engaging key stakeholders to advocate for the surgical patient. As an example, the G4 Alliance, comprising over 80 organizations in more than 160 countries, has played a strong role in advocacy, policy implementation, and strategic planning for global surgery as outlined in their forthcoming Global Action Plan for Surgery.36 Another example includes the Optimal Resources for Children’s Surgery, developed by the Global Initiative for Children’s Surgery, a collaboration of more than 270 pediatric health care professionals and academics from 44 countries (75% from LMICs).37,38 The document outlined detailed standards of children’s surgical care at all levels of health care facilities, with the objective to inform national surgical plans in LMICs.

### Conclusion

We identified key resources as relevant to practitioners from various specialties and stages of their careers and listed the top 10 resources relevant to individuals across all groups. We hope that this list will serve as a starting point for individuals interested in global surgery and aid in promoting global surgery awareness, scholarship, and collaboration toward a world in which safe surgical and anesthesia care is a reality for all.

**Competing interests:** None declared.

### References


Peer Reviewed

Received: January 30, 2020, Accepted: August 5, 2020


© Nyikuri et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are properly cited. To view a copy of the license, visit http://creativecommons.org/licenses/by/4.0/. When linking to this article, please use the following permanent link: https://doi.org/10.9745/GHSP-D-20-00050

Global Health: Science and Practice 2020 | Volume 8 | Number 3
In the article “Coping with COVID-19: learning from past pandemics to avoid pitfalls and panic” by Daniel T. Halperin, which appeared in the June 2020 issue (Volume 8, Issue 2), on page 1, the fourth sentence under the heading “Confusion and Panic Returns With This Pandemic,” was changed to “Under such circumstances, fear is understandable and can help motivate behavior change. However, when fear becomes irrational or leads to panic, it often results in poor decision making and other unintended consequences.” to avoid the incorrect implication that irrational fear can have a positive effect on motivating behavior change.

The article has been corrected accordingly.