EDITORIALS

Learning From Neighbors

We can learn valuable lessons from program efforts that at first glance may seem to be far removed from our own work.

Stephen Hodgins

Glob Health Sci Pract. 2020;8(4):613
https://doi.org/10.9745/GHSP-D-20-00639

COMMENTARIES

Go Where the Virus Is: An HIV Micro-epidemic Control Approach to Stop HIV Transmission

Essentially all HIV transmission is from people living with HIV who are not virally suppressed. An HIV micro-epidemic control approach that differentiates treatment support and prevention services for people living with HIV and their network members according to viral burden could optimize the impact of epidemic control efforts.

Michael M. Cassell, Rose Wilcher, Reshmie A. Ramautarsing, Nittaya Phanuphak, Timothy D. Mastro

https://doi.org/10.9745/GHSP-D-19-00418

Capturing Acquired Wisdom, Enabling Healthful Aging, and Building Multinational Partnerships Through Senior Global Health Mentorship

The undeniable benefit of mentorship by experience senior mentors can meaningfully increase the breadth of their experience and contributions to society as well as address the dire inequality in global health. This model captures wisdom lost to retirement, enables opportunities for purposeful lifespan, underpins sustainable health care systems, and has the potential for building multinational partnerships.

C. Norman Coleman, John E. Wong, Eugenia Wendling, Mary Gospodarowicz, Donna O’Brien, Taofeeq Abdullahh Ige, Simeon Chinedu Aruah, David A. Pistenmaa, Ugo Amaldi, Onyi-Onyinye Balogun, Harmar D. Brereton, Silvia Formenti, Kristen Schroeder, Nelson Chao, Surbhi Grover, Stephen M. Hahn, James Metz, Lawrence Roth, Manjit Dosanjha

https://doi.org/10.9745/GHSP-D-20-00108

ORIGINAL ARTICLES


Internally displaced persons fleeing violent conflict represent a neglected population with heightened vulnerability to pandemic COVID-19. We provide a rare snapshot of the overwhelming challenges faced by internally displaced persons in Eastern Democratic Republic of the Congo as they brace for COVID-19.

Kasereka M. Claude, Muyisa Sahika Serge, Kahindo Kahatane Alexis, Michael T. Hawkes

https://doi.org/10.9745/GHSP-D-20-00108
Meeting the Global Target in Reproductive, Maternal, Newborn, and Child Health Care Services in Low- and Middle-Income Countries

What progress has been achieved toward reproductive, maternal, newborn, and child health service related Sustainable Development Goals? Analyzing data to estimate coverage of these indicators, we observed that acceleration is needed in coordinated global efforts and government policies to ensure universal access to RMNCH care services by 2030.


https://doi.org/10.9745/GHSP-D-20-00097

Contraceptive Method Mix: Updates and Implications

Trends in contraceptive method mix show that dominance of 1 method in the mix remains very common, though countries and regions throughout the world are diverse as to which method is dominant. Our analysis argues for continued concerted efforts of programs to increase contraceptive method choice.

Jane T. Bertrand, John Ross, Tara M. Sullivan, Karen Hardee, James D. Shelton

https://doi.org/10.9745/GHSP-D-20-00229

Health Care Worker Preferences and Perspectives on Doses per Container for 2 Lyophilized Vaccines in Senegal, Vietnam, and Zambia

When providing immunization services, health care workers balance the mandate of achieving high coverage with limiting vaccine wastage. Workers in 3 countries said that containers with fewer vaccine doses for measles and BCG would enable them to immunize all children who present, while reducing concerns about wasting vaccine.

Natasha Kanagat, Kirstin Krudwig, Karen A. Wilkins, Sydneey Kaweme, Guissimon Phiri, Frances D. Mwansa, Mercy Mvundura, Joanie Robertson, Debra Kristensen, Abdoulayaye Gueye, Sang D. Dao, Pham Q. Thai, Huyen T. Nguyen, Thang C. Tran

https://doi.org/10.9745/GHSP-D-20-00112

Remote Mentorship Using Video Conferencing as an Effective Tool to Strengthen Laboratory Quality Management in Clinical Laboratories: Lessons From Cambodia

This program to strengthen laboratory quality management systems in Cambodia demonstrated significant improvements in conformity to ISO 15189 standards in participating laboratories, correlating with laboratory participation time in video conference training activities led by quality improvement mentors over the program implementation period.

Grant Donovan, Siew Kim Ong, Sophanna Song, Nayah Ndefru, Chhayheng Leang, Sophat Sek, Patricia Sadate-Ngatchou, Lucy A. Perrone

https://doi.org/10.9745/GHSP-D-20-00128
Using Community Health Workers and a Smartphone Application to Improve Diabetes Control in Rural Guatemala

A smartphone application providing algorithmic clinical decision support enabled community health workers to improve diabetes control for a group of patients in rural Guatemala. This approach enables task sharing with physicians and other advanced practitioners for chronic disease care, which is particularly important in low-resource settings.

Sean Duffy, Derek Norton, Mark Kelly, Alejandro Chavez, Rafael Tun, Mariana Niño de Guzmán Ramírez, Guanhua Chen, Paul Wise, Jim Svenson

https://doi.org/10.9745/GHSP-D-20-00076

Behavioral Insights Into Micronutrient Powder Use for Childhood Anemia in Arequipa, Peru

Health care provider-caregiver interactions and caregivers’ shifting emotional states between intention formation and use affected their adherence to a government-provided micronutrient powder (MNP) intervention to prevent childhood anemia. In counseling directed to caregivers, we suggest providers offer clear messaging on MNP impact and planning for challenges during MNP use.

Jessica D. Brewer, Julianna Shinnick, Karina Román, Maria P. Santos, Valerie A. Paz-Soldan, Alison M. Buttenheim

https://doi.org/10.9745/GHSP-D-20-00078

A Cluster-Randomized Trial to Test Sharing Histories as a Training Method for Community Health Workers in Peru

Women naturally communicate using life narratives. Through systematic recall and sharing memories of their own childbearing and child rearing experiences, community health workers (CHWs) become engaged and empowered to change their own and other mothers’ health behaviors. Training CHWs with Sharing Histories can improve their capabilities as change agents for better child health.

Laura C. Altobelli, José Cabrejos-Pita, Mary Penny, Stan Becker

https://doi.org/10.9745/GHSP-D-19-00332

A Rapid Cost Modeling Tool for Evaluating and Improving Public Health Supply Chain Designs

The Rapid Supply Chain Modeling Tool enables health system leaders to quickly estimate and compare the cost impact of potential supply chain design improvements in situations where time and budget do not allow for more in-depth modeling approaches.

Michael Krautmann, Mariam Zameer, Dorothy Thomas, Nora Phillips-White, Ana Costache, Pascale R. Leroueil

https://doi.org/10.9745/GHSP-D-20-00227
Lessons Learned From Implementing Prospective, Multicountry Mixed-Methods Evaluations for Gavi and the Global Fund

Lessons learned from implementing evaluations for Gavi, the Vaccine Alliance and the Global Fund for AIDS, Tuberculosis, and Malaria can help inform the design and implementation of ongoing or future evaluations of complex interventions. We share 5 lessons distilled from over 7 years of experience implementing evaluations in 7 countries.

Emily Carnahan, Nikki Gurley, Gilbert Asiimwe, Baltazar Chilundo, Herbert C. Duber, Adama Faye, Carol Kamya, Godefroid Mpanya, Shakilah Nagasha, David Phillips, Nicole Salisbury, Jessica Shearer, Katharine Shelley, for the Gavi Full Country Evaluations Consortium; and Global Fund Prospective Country Evaluation Consortium

https://doi.org/10.9745/GHSP-D-20-00126

Effects of a Community-Based Program on Voluntary Modern Contraceptive Uptake Among Young First-Time Parents in Cross River State, Nigeria

Among young first-time mothers, participation in a comprehensive, community-based program led to a 3-fold increase in voluntary modern contraceptive use and other positive changes. These results demonstrate the importance of investing in interventions for this youth population that ideally address a range of priorities across the first-time parent lifestage.

Gwendolyn Morgan, Anjala Kanesathasan, Akinsewa Akiode

https://doi.org/10.9745/GHSP-D-20-00111

REVIEWS

Routine Family Planning Data in the Low- and Middle-Income Country Context: A Synthesis of Findings From 17 Small Research Grants

A review of 5 years of small grant-funded research highlighted overarching barriers to and opportunities for using family planning data in routine health information systems in low- and middle-income countries. We report on factors affecting data quality, analysis, and use, and suggest strategies to improve routine family planning data.

Bridgit Adamou, Janine Barden-O’Fallon, Katie Williams, Amani Selim

https://doi.org/10.9745/GHSP-D-20-00122

Effectiveness of mHealth Interventions for Improving Contraceptive Use in Low- and Middle-Income Countries: A Systematic Review

Do mHealth interventions help reduce unmet contraceptive needs in low- and middle-income countries by attempting to increase the uptake of modern contraceptive methods? Which mHealth features and behavior change communication components were used in these mHealth interventions? This review aimed to answer these questions and assess the impact of these interventions on contraceptive uptake outcomes.

Banyar Aung, Jason W. Mitchell, Kathryn L. Braun

https://doi.org/10.9745/GHSP-D-20-00069
A Practical Guide to Using Time-and-Motion Methods to Monitor Compliance With Hand Hygiene Guidelines: Experience From Tanzanian Labor Wards

Understanding hand hygiene behaviors is critical in hospitals. We developed the HANDS at birth tool—and provide information on its design and implementation—to capture the complex patterns of health care workers’ hand hygiene including hand rubbing/washing, glove use, and recontamination.

Giorgia Gon, Said M. Ali, Robert Aunger, Oona M. Campbell, Micheál de Barra, Marijn de Bruin, Mohammed Juma, Stephen Nash, Amour Tajo, Johanna Westbrook, Susannah Woodd, Wendy J. Graham

https://doi.org/10.9745/10.9745/GHSP-D-20-00221

Implementation of a Pediatric Early Warning Score to Improve Communication and Nursing Empowerment in a Rural District Hospital in Rwanda

Implementation of the Pediatric Early Warning Score for Resource-Limited Settings tool improved nurses’ competency and confidence in their triage capabilities. This tool has the potential to improve patient outcomes. However, staff turnover and limited physician buy-in were barriers to sustainability of the tool in low-resource settings.

Shela Sridhar, Alexis Schmid, Francois Biziyaremye, Samantha Hodge, Ngamika Patient, Kim Wilson

https://doi.org/10.9745/GHSP-D-20-00075

Juntos: A Support Program for Families Impacted by Congenital Zika Syndrome in Brazil

Development of Juntos, a community-based family support program for caregivers of children with congenital Zika syndrome, contained innovative approaches such as using mothers as facilitators, fast-track learning, and de-isolation of families affected by Zika.

Antony Duttine, Tracey Smythe, Miriam Ribeiro Calheiros de Sa, Silvia Ferrite, Maria Elisabeth Moreira, Hannah Kuper

https://doi.org/10.9745/GHSP-D-20-00018
TECHNICAL NOTES

Improving Hospital Oxygen Systems for COVID-19 in Low-Resource Settings: Lessons From the Field

Hospitals in low- and middle-income countries urgently need to improve their oxygen systems for COVID-19 and other health emergencies. We share practical tips to improve pulse oximetry and oxygen use, support biomedical engineers to optimize existing oxygen supplies, and expand existing oxygen systems with robust equipment and smart design.

Hamish R. Graham, Sheillah M. Bagayana, Ayobami A. Bakare, Bernard O. Olayo, Stefan S. Peterson, Trevor Duke, Adegoke G. Falade

Glob Health Sci Pract. 2020;8(4):858–862
https://doi.org/10.9745/GHSP-D-20-00224

CORRECTIONS

Corrigendum: Parmaksiz K et al., What Makes a National Pharmaceutical Track and Trace Succeed? Lessons From Turkey

Glob Health Sci Pract. 2020;8(4):863
https://doi.org/10.9745/GHSP-D-20-00587
The article by Duttine and colleagues in this issue of Global Health: Science and Practice describes efforts recently undertaken in Brazil to develop and test an approach to supporting parents of young children affected by Zika. For the vast majority of GHSP readers, we suspect that this topic is not directly relevant to their work. Nevertheless, we believe we can learn valuable lessons from program efforts that may seem, at first glance, to be far removed from our own work.

The authors describe their process in the development and initial testing of a community-based support program for parents. Their starting points were: (1) a clearly defined need, and (2) a potentially relevant model. The need was for effective, formalized support of caregivers of community-living children with mild to moderate Zika-related impairments. The model was a caregiver education and support program first developed in Bangladesh and subsequently adopted in many other countries: Getting to Know Cerebral Palsy.

With this identified gap and a potential solution, the next step was a needs assessment, which included a review of relevant literature and consultation with stakeholders, both experts and parents of affected children. On the basis of this input, the developers then crafted an initial version of the program design and materials (adapted to Zika and to a Brazilian cultural setting) and an associated theory of change (a theory of how they thought the intervention would work).

The program developers then tested the program in 2 diverse sites, going through 2 iterative rounds of piloting to assess relevance, usefulness, and feasibility. Adaptation and refinement of the program was done not only at the end of each of these rounds; the developers also pursued an intentional approach of fast-track learning—eliciting feedback on an ongoing basis from participants and facilitators and making real-time changes, as necessary, in content, approach, and logistical arrangements.

As they explain, not only was the program itself iteratively adapted and modified, so too was the theory of change. In the course of trying out the intervention, the investigators formed a better sense of how it worked and, in turn, made revisions to their theory of change.

The authors modeled learning from neighbors, drawing on an approach developed in another context for a somewhat different need (Bangladeshi families with children affected by cerebral palsy) and adapting it to their setting to address their specific problem. We encourage you—our readers—also to learn from neighbors, drawing from examples like this of how to address a public health problem by:

- Listening to stakeholders
- Being flexible and prepared to revisit assumptions and early design choices
- Building learning and adaptation into your routine ways of doing business

Competing interests: None declared.

REFERENCE

See related article by Duttine et al.

Learning From Neighbors

Stephen Hodgins

Global Health: Science and Practice 2020 | Volume 8 | Number 4
Go Where the Virus Is: An HIV Micro-epidemic Control Approach to Stop HIV Transmission

Michael M. Cassell, Rose Wilcher, Reshmie A. Ramautarsing, Nittaya Phanuphak, Timothy D. Mastro

Key Messages

- Essentially all HIV transmission is from people living with HIV (PLHIV) who do not know their infection status or have not yet achieved viral suppression, making support for these individuals and their risk contacts a priority for treatment and prevention efforts.

- Proven approaches exist to reduce viral burden and interrupt HIV transmission from PLHIV who are not yet virally suppressed, but these approaches must be implemented with enhanced focus and scale to maximize benefit.

- Improved diagnostic approaches offer new opportunities to increase public health impact by prioritizing support for unserved or underserved individuals with the greatest viral burdens and among members of their risk networks.

- Policy makers should pursue the implementation and evaluation of diagnostic approaches that can focus services among individuals and networks with the greatest viral burdens.

- Program managers should treat viral burden as a primary consideration in the provision of differentiated HIV services, applying an HIV micro-epidemic control framework to prioritize and tailor services for PLHIV and their risk contacts along a continuum of progression to viral suppression.

INTRODUCTION

Globally, a growing majority (59%) of an estimated 38 million people living with HIV (PLHIV) know their HIV status and have achieved HIV viral suppression by adhering to antiretroviral therapy (ART). Individuals who achieve sustained viral suppression and undetectable levels of circulating virus through good adherence to ART live long, healthy lives and will not transmit HIV through sexual contact. The evidence that people who have achieved undetectable viral loads will not transmit HIV sexually—that “undetectable equals unttransmittable” (U=U)—underscores the prevention benefits of treatment and the rationale for the global call to achieve near-universal access to ART and viral suppression among PLHIV.

Conversely, HIV viral burden (viremia), generally measured by plasma viral load (HIV RNA copies/mL) assays, is the primary predictor of HIV-related disease progression, morbidity, mortality, and ongoing transmission. Essentially all HIV transmission originates from a shrinking minority of PLHIV globally (41%) who do not know their HIV infection status or have not yet achieved viral suppression, making support for these individuals and their risk contacts a priority for treatment and prevention efforts. Studies have identified a dose-response relationship in which each 10-fold increase in HIV plasma viral load results in an increased relative risk of HIV transmission of 2.5 to 2.9 per sexual contact. Emerging evidence suggests that even under conditions of near-universal HIV treatment coverage, high viremia and high levels of risk behavior among unserved or underserved PLHIV can sustain epidemic HIV transmission. In a recent U.S. study of HIV patients in care with a detectable viral load, only a small proportion of PLHIV reported concurrent sexual transmission risk behaviors, but most of the individuals in this group had considerably elevated viral loads, increasing the probability of transmission. The study found that viral loads were likely to be lower among those with a detectable viral load who reported always using condoms.

High viral burden associated with acute HIV infection (AHI) is a particular concern. Acute infection is characterized by a 2-4-week period of exceptionally high viremia as HIV replicates rapidly in the body before a person’s immune system mounts a response and reduces the level of circulating virus to much lower—but typically not undetectable—levels for a period of months to years. Although only a small proportion...
of all PLHIV will be in this brief AHI phase at any given time, per-sex-act transmission probabilities are considerably higher during periods of acute as compared to chronic HIV infection.9,16–18

In key populations engaged in frequent behavioral risks, up to an estimated 50% of all HIV transmission occurs from individuals during AHI when viremia is very high prior to the development of an immune response including anti-HIV antibodies (Ab) that yield reactivity on third-generation Ab assays.14,17,19–24 The provision of ART during AHI and of HIV pre-exposure prophylaxis (PrEP) to the risk-network contacts of acutely infected individuals could prevent a substantial proportion of ongoing HIV transmission. An analysis in Thailand suggested that early diagnosis and treatment during AHI among men who have sex with men could avert 89% of all new infections in this population.25

Approaches that differentiate service delivery to better address the preferences and needs of unserved and underserved individuals have been identified as a priority to close outstanding gaps in access to HIV prevention and treatment.26 In implementing differentiated services, it is increasingly clear that a focus on individuals and networks with the greatest viral burdens has strategic benefit. For example, programs typically transition individuals who are receiving HIV treatment and are identified through routine viral load testing as virally suppressed to options for less frequent clinical follow-up and multimonth dispensing of their antiretroviral medications. This differentiation offers additional convenience to patients and frees up resources and provider time to focus support on virally unsuppressed individuals with greater adherence, clinical, social support, and other needs.

Nevertheless, the resources and technologies needed to activate a more comprehensive differentiation of support based on viral burden historically have been limited. With the advent of expanded access to viral load testing and options to screen for AHI, opportunities now exist to prioritize support for individuals and in risk networks with the greatest viral burdens. This prioritization can help interrupt epidemic HIV transmission associated with AHI through early diagnosis, HIV treatment, and provision of PrEP and other proven prevention approaches to risk contacts. Because HIV morbidity, mortality, and transmission risk are most closely associated with viral burden, this enhanced focus can guide the allocation of limited resources to maximize the impact of prevention and treatment efforts.

**ENVISIONING A MICRO-EPIDEMIC CONTROL APPROACH THAT DIFFERENTIATES SUPPORT BASED ON VIRAL BURDEN**

We propose an HIV micro-epidemic control framework to characterize these opportunities to accelerate impact, with a primary focus on addressing the differentiated service preferences and needs of individuals who are not yet virally suppressed, as well as the members of their risk networks. This framework aims to organize and integrate both new and existing approaches to tailor support for PLHIV and their risk contacts based on progression to sustained viral suppression. By profiling the characteristics of
those who face challenges in achieving viral suppression—such as barriers to diagnosis, to treatment initiation and retention post-diagnosis, and to viral load testing and suppression—programs can introduce solutions that both help these individuals and that remove barriers for others.

The approach features variable treatment and prevention services and service intensity grouped according to 4 different stages along this continuum of progression to viral suppression (Figure 1). The model also affords program managers with opportunities to prioritize program efforts based on regional, national, and subnational variations in progress with respect to the expansion of HIV prevention, testing, treatment, and viral suppression coverage.

1. PLHIV Who Are on Treatment and Virally Suppressed

Although progress varies by region, an estimated 59% of PLHIV globally are already receiving HIV treatment and are virally suppressed. Importantly, per “U=U,” these individuals will not transmit HIV to their sexual partners. PLHIV who are found to be virally suppressed through routine viral load testing are good candidates for multimonth dispensing, which will reduce the need for more frequent clinic visits, a benefit for both PLHIV and clinic staff.

Partner notification services, also known as index testing, are recommended by the World Health Organization as a safe, effective strategy to accelerate HIV epidemic control by asking PLHIV to list and refer their sexual and injecting partners and biological children to HIV testing services on a voluntary basis. Offering index testing at least once to these individuals can help link members of their networks who may previously have been exposed to HIV to relevant testing, prevention, and treatment services. Uninfected network members will not acquire HIV infection from sexual contact with PLHIV who have undetectable viral loads, but those who continue to be at elevated HIV infection risk from other contacts can be offered PrEP and other HIV prevention services, including condom education and access. Routine viral load testing is critical to monitor and sustain viral suppression among PLHIV who have previously achieved suppression. To facilitate viral load testing access and the provision of efficient retention and adherence support to PLHIV with a stable treatment status, programs can implement virtual online- and telephone-based support with patient consent and appropriate measures in place to

---

**FIGURE 1.** An HIV Micro-epidemic Control Model Aims to Prioritize and Focus Treatment and Prevention Efforts Where They Can Have the Greatest Impacts: Among a Shrinking Proportion of Individuals and Risk Networks With the Greatest Viral Burdens

---

Abbreviations: AHI, acute HIV infection; ART, antiretroviral therapy; HIV+, HIV-positive; HIV-, HIV-negative; PLHIV, people living with HIV; PrEP, pre-exposure prophylaxis.
ensure the security and confidentiality of patient information. The use of point-of-care viral load testing technologies may also reduce testing turn-around times and bring added convenience to patients and providers.

2. PLHIV Who Are Diagnosed but not Virally Suppressed

A substantial proportion of individuals who have previously received an HIV diagnosis have either not yet initiated ART or have not achieved viral suppression. Individuals in this group can be further divided into 3 categories: (1) those who have never been linked to ART; (2) those who have initiated ART but have not yet achieved viral suppression or have been lost to follow-up and stopped ART; and, (3) those who are sustained on treatment but are showing signs of breakthrough viremia or treatment failure. For individuals who have never initiated treatment or been lost to follow-up, programs can initiate outreach campaigns through clinical or community staff to engage or reengage previously diagnosed individuals. These campaigns can promote “U=U” messaging, the benefits of new dolutegravir-based treatment regimens, and convenient and confidential options for PLHIV to access same-day HIV treatment. For individuals who are receiving treatment but have not achieved viral suppression, providers can offer additional personalized adherence counseling and support. Immediate support should be provided to individuals with a viral load test indicating an unsuppressed viral load to help to identify and address root causes of adherence or treatment failure that require regimen switching.

In the process of providing personalized treatment support to PLHIV, a range of voluntary index testing options can be offered to encourage referrals of their risk contacts to prevention services such as PrEP and condoms until they have achieved viral suppression, or to HIV treatment services, as needed. It may also be useful for programs to routinely monitor the sociodemographic and risk characteristics of individuals who do not initiate treatment, fall out of care, or do not achieve viral suppression, to assess how these individuals differ from those who engage in treatment and sustain good treatment outcomes. By generating profiles of the characteristics of individuals more likely to face treatment challenges, programs can apply these to prevent loss and other adverse outcomes, helping to accelerate and sustain progression to viral suppression. Some programs are applying machine learning algorithms to automate this process of preventive prioritization to enhance care.

3. Undiagnosed PLHIV

PLHIV who have not yet received a diagnosis can be similarly offered tailored support to maximize individual treatment and population-level prevention benefits. An expanded range of options for accessing HIV testing services, including HIV self-testing options with dispensing through pharmacies and peer networks, can help to close gaps in diagnosis among PLHIV who might otherwise not otherwise access diagnostic or other services. Testing services also can be tailored to focus on key populations facing the greatest HIV infection risks and to engage the risk network members of PLHIV who are not virally suppressed. Incorporating AHI screening into these targeted testing approaches can improve the detection of AHI among individuals who might otherwise have remained undiagnosed. While offering voluntary index testing and risk contact referral services to individuals who are newly diagnosed or who have recent HIV infection as identified through recency testing, programs can also assess the differentiating sociodemographic and risk profiles of these individuals. These profiles can be applied to further enhance the focus of HIV testing efforts by bringing testing to individuals with similar profiles and by engaging peer mobilizers with similar characteristics to make testing referrals and distribute HIV self-testing kits in their networks.

4. PLHIV With AHI

Expanding screening for AHI among key populations and other individuals facing elevated HIV infection risks can help realize the largely untapped treatment and prevention benefits of identifying, treating, and index testing individuals with AHI. To maximize efficiency, AHI screening can be prioritized for the risk network members of individuals identified with AHI, those with recent HIV infections, and newly identified PLHIV. Screening can also be focused on key populations reporting recent behavioral risks, as well as among those with other sexually transmitted infections. Upon diagnosis, all PLHIV can be offered an accelerated path to viral suppression with same-day treatment initiation. Programs can assess the differentiated characteristics of newly diagnosed, recently infected, and acutely infected PLHIV to further optimize the relevance and focus of HIV testing efforts.
The HIV micro-epidemic control approach aims to enhance prevention benefit by focusing interventions in the risk networks in which active HIV transmission is occurring.

The brief duration of acute infection and cost of testing poses barriers to diagnosis.

While the treatment and treatment-as-prevention benefits of prioritizing support according to viral burden may have evident advantages, the HIV micro-epidemic control approach also aims to enhance prevention benefit by focusing services in the risk networks in which active HIV transmission is occurring. The majority of risk contacts of PLHIV who are undiagnosed will be uninfected but at high risk of acquiring HIV infection, making linkages of these individuals to prevention services a priority. These focused prevention efforts should employ a combination of evidence-based prevention strategies relevant to the specific preferences and needs of the populations being served. These strategies include, but are not limited to, harm reduction programming for people who inject drugs, support for correct and consistent condom use, and expanded access to PrEP. To maximize uptake, services should be implemented in a friendly manner that is welcoming and convenient to clients and is responsive to their feedback.

For all risk contacts of PLHIV who are not yet virally suppressed, PrEP is a critical, evidence-based, and likely short-duration priority. Making PrEP—and, as relevant, nonoccupational HIV post-exposure prophylaxis—offers routine for the contacts of PLHIV as part of index testing affords enormous opportunities to focus PrEP where it can have the greatest prevention impact. In circumstances where partners and other risk contacts face no other substantial HIV infection risks, these individuals can safely discontinue PrEP once the PLHIV index client has achieved viral suppression. In addition, the scale up of PrEP as part of index testing services may serve to normalize PrEP and expand availability across a wider array of settings, helping to accelerate historically limited progress towards the achievement of global PrEP targets and removing barriers to access among men who have sex with men and individuals who may prefer not to disclose their status as key populations.

Stigma, discrimination, violence, and other structural factors impose considerable barriers to service uptake, particularly among PLHIV and key populations living in criminalizing environments. Therefore, the micro-epidemic control approach should be implemented in conjunction with broader efforts to address these structural factors. By identifying the characteristics of individuals for whom structural factors serve as particular barriers to health and well-being, a micro-epidemic control framework may help bring additional focus to structural interventions and facilitate advocacy and partnerships with individuals and communities to develop and implement voluntary, safe, equitable, and preferred policy and program solutions.

THE CHALLENGE OF ACUTE HIV INFECTION

The proposed micro-epidemic control approach emphasizes diagnosis of and intervention during AHI in light of the substantial role that AHI plays in epidemic transmission of HIV. Most current national HIV testing algorithms rely on antibody-based serological testing that cannot detect AHI. As a result, these approaches misdiagnose potential core transmitters as HIV-uninfected and miss critical opportunities to maximize the prevention benefits of HIV treatment. Affordable, accurate, and scalable solutions to diagnose AHI have remained elusive.

The brief duration of AHI poses a major challenge to diagnosis. Detection of AHI depends on infected individuals having a blood test during the short AHI period and then establishing the presence of HIV RNA or p24 antigen (part of the virus) (Figure 2). Individuals facing high infection risks would need to seek HIV testing with HIV RNA or p24 technologies on a frequent basis to increase the likelihood of detecting an infection during the acute period.

Cost is also an issue. Point-of-care platforms for detection of HIV RNA such as Alere Q (Abbott Laboratories) and GeneXpert (Cepheid) are now available but are generally perceived as expensive (US$17–24). Fourth-generation point-of-care rapid HIV tests, such as the Alere HIV Combo kit (Abbott Laboratories), detect both p24 antigen and HIV antibodies within 20 minutes at a lower cost (US$2–4) and can be substituted as the first, sensitive screening test in a national HIV testing algorithm for diagnosis. However, these fourth-generation assays have much lower sensitivity to detect AHI than HIV RNA assays.

Nevertheless, a clinical trial of PrEP in Uganda, South Africa, and Zimbabwe found that 28% of infections missed by current third-generation rapid diagnostic tests would have been identified with the use of Alere HIV Combo, suggesting some advantages of using a fourth-generation test over standard antibody testing. Investigators in San Francisco found more promising results, with the Alere Determine (Abbott Laboratories) point-of-care fourth-generation antigen-antibody combo rapid test detecting about 54% of the acute cases detected through laboratory RNA testing. In Thailand, the Alere HIVCombo kit detected 37 of 50 (74%) individuals with HIV RNA confirmed
acute HIV infection. These limited results suggest promising performance of Alere HIV Combo in a facility-based setting but require broader evaluation in diverse settings and populations.42,43

AHI is sometimes accompanied by transient clinical flu-like and other symptoms, including rash, fever, sore throat, fatigue, muscle/joint aches, oral and genital ulcers, diarrhea, and swollen lymph nodes.44,45 However, these symptoms and signs are neither sensitive nor specific for AHI. Inquiring about the presence of these symptoms and recent risk behaviors may suggest opportunities to screen for AHI with an RNA assay if available or at least a fourth-generation test.46–48 Sensitizing populations at risk to AHI signs and symptoms, benefits of early detection and treatment, and potential advantages of fourth-generation diagnostics may also facilitate improved AHI diagnosis and treatment and mitigate HIV resistance risks associated with PrEP continuation among individuals who may have received false-negative third-generation HIV testing results.

To overcome some of the logistical and resource-related barriers to the expansion of AHI screening, diagnosis, and intervention, we propose 2 potential solutions to encourage further consideration of both these and other context-relevant approaches: (1) conducting pooled HIV-1 RNA polymerase chain reaction (PCR) on dried blood spot samples; and (2) expanding the use of fourth-generation point-of-care rapid HIV testing, leveraging recency testing data where possible to help focus AHI screening in networks in which ongoing HIV transmission is occurring.

**POOLED PCR TO FACILITATE DETECTION OF AND INTERVENTION DURING AHI**

The “gold standard” for detection of AHI is molecular testing, specifically HIV-1 RNA PCR. This approach is considered the standard of care to facilitate early infant diagnosis among children born to HIV-infected mothers. However, PCR is relatively expensive.17 To extend PCR testing efficiently to all individuals facing elevated HIV infection risks but who have nonreactive serological testing results as part of a targeted HIV testing strategy, samples can be “pooled” such that qualitative PCR is run on a batch that combines like sample types sourced from different individuals. Individual results are confirmed as negative for negative pools. For reactive pools, each sample in the pool is then tested with quantitative PCR viral load independently to identify and rapidly
intervene with individuals with reactive results. The Thai Red Cross AIDS Research Center has been applying a pooling approach with plasma samples as a cost-efficient strategy to identify and treat individuals with AHI who might otherwise not receive a diagnosis using serological testing. Nevertheless, separation, storing, and transfer of plasma can pose logistical challenges in resource-limited settings and incur additional costs. Pooled PCR testing may also be possible on point-of-care viral load platforms as these become more affordable and widely available, realizing additional benefits in terms of efficiency, convenience, and early detection and intervention.

In resource-limited settings, collecting dried blood spot samples for pooled PCR is a promising approach to circumvent some of the cost barriers associated with point-of-care platforms and the logistical challenges associated with separation, storing, and transfer of plasma samples. Whole blood spots can be collected with a finger prick and can be stored and shipped with relative ease. Recalibration of the PCR is necessary because of a degradation of viral RNA in dried blood that may result in a 2-log reduction in assay sensitivity and because of the potential presence of viral DNA, which may partially compensate for this loss in sensitivity. At least 1 study has demonstrated the accuracy of doing pooled PCR on dried blood spot samples for early infant diagnosis and documented a laboratory cost savings of 65% associated with pooling. Other studies have demonstrated feasibility to diagnose AHI and to reduce the costs of ART monitoring in resource-limited settings with pooled PCR on dried blood spot samples.

An illustrative depiction of this diagnostic algorithm incorporating AHI screening based on dried blood spot sample pooled PCR testing is provided in Figure 3.

To maximize the benefits of screening for AHI, the time from sample collection to case identification and intervention must be minimized. A pooling approach is more practical in high volume settings in which batches can be run every day. While the time needed to process samples will vary according to the proximity and availability of laboratory infrastructure, we anticipate many programs being able to provide results in 1–2 days given the predominantly urban concentration of laboratory resources and of HIV key population risks in many country settings. The expansion of point-of-care viral load testing can further reduce turnaround times, facilitating earlier action.

The routine collection of dried blood spot samples to detect and intervene during AHI also presents some additional value-added opportunities to improve HIV service delivery that are worthy of consideration. Drug-resistance genotyping can be performed on samples from all individuals who are confirmed through either molecular or serological testing to have HIV infection. In addition, these samples could be used to conduct phylogenetic analyses of potential HIV transmission clusters to enhance the focus of targeted testing and index testing implementation.

### COMBINING AHI SCREENING WITH RECENCY TESTING TO FACILITATE EARLIER HIV DIAGNOSIS AND INTERVENTION

Rapid HIV recency assays, such as the Asanté HIV-1 Rapid Recency® Assay (Sedia Biosciences) and the Maxim Swift HIV Recent Infection Assay (Maxim Biomedical Inc.), were developed to help identify individuals who have become HIV infected within the past year—on average in the past 6 months—to help estimate HIV incidence and improve the focus of programming in settings, populations, and networks in which incident infections continue to occur. Rapid recency point-of-care antibody-based assays differentiate between recent HIV infection—when the antibody response is immature, as reflected by low “avidity” or binding strength of the antibody—and long-term infections in which a mature antibody response is measured by strong antibody avidity. The assays can yield “false-recent” results among individuals who naturally control HIV well (low virus=low antibody) or are receiving ART, so a recent infection result is usually confirmed using a recent infection testing algorithm in which a viral load test is conducted with results of ≥ 1,000 copies/mL confirming recent infection.

Rapid recency assays only measure antibody avidity after HIV seroconversion; they do not detect HIV RNA or p24 antigen and therefore are unable to detect AHI. In typical use, they are only offered to individuals who have been confirmed to have HIV infection with a standard antibody-based national HIV testing algorithm. Recency assays are also pending review for diagnostic purposes by the World Health Organization and are currently only approved for research use by the U.S. Food and Drug Administration. The World Health Organization has endorsed the use of recency assays for surveillance purposes but has not yet made a determination regarding program...
or individual-level benefits pending further evidence.54

That said, the rapid recency algorithm has demonstrated capacity to identify individuals who became HIV infected within the past year, and the U.S. President’s Emergency Plan for AIDS Relief has identified the scale up of recency testing as a “minimum standard” for HIV program implementation in an expanding set of countries receiving U.S. government support.59 While there currently is no rationale to offer differentiated counseling or clinical HIV treatment support to PLHIV with recent versus long-term HIV infection, it is more likely that persons with recent infections are part of ongoing transmission networks. Individuals with recent infection were recently acutely infected and were recently in risk contact with at least 1 other PLHIV who was not virally suppressed. Therefore, targeting testing among the contacts of recently infected individuals could improve the capacity of programs to detect and treat previously undiagnosed individuals while focusing prevention services among individuals facing the greatest infection risks. Moreover, conducting AHI screening among the network contacts of recently infected individuals, as well as targeting AHI screening among individuals with similar risk and sociodemographic profiles to those with recent infections, could increase capacity to detect, treat, and prevent transmission during AHI.

To leverage recency testing data to help focus AHI screening as part of an HIV micro-epidemic control model, programs would need to adopt

Testing the contacts of recently infected individuals could improve a program’s capacity to detect and treat undiagnosed individuals.
strategies to: (1) integrate AHI screening into practice; (2) secure client informed consent for recency testing and the confidential use of those results; (3) support confidential profiling of individuals with recent and acute infections; (4) target testing with AHI screening in populations, settings, and networks aligned to these profiles; and (5) prioritize index testing with AHI screening among the contacts of recently and acutely infected individuals.

An illustrative workflow for AHI screening supplemented by recency testing to improve focus is shown in Figure 4. AHI screening could be conducted using a pooled PCR approach on dried blood spot samples as previously described. However, in this instance, we outline an approach in which members of key and priority populations could be offered screening for AHI through combined use of a sensitive fourth-generation rapid diagnostic test like the Alere HIV Combo, as well as a risk- and symptom-based verbal screening tool. A potentially useful example of a tool validated using AHI data from the Amsterdam Cohort Studies among men who have sex with men consisted of a self-administered weighted survey inquiring about the presence of 4 current symptoms (fever, lymphadenopathy, oral thrush, and weight loss), and 3 risk factors in the past 6 months (receptive condomless anal intercourse, more than 5 sexual partners, and gonorrhea).

Prioritizing AHI screening as part of index testing for the risk network contacts of recently infected individuals, as well as eventually for the contacts of individuals with AHI as these are identified, could then help programs increase the likelihood of detecting people with AHI. Case profiles can be developed to describe actionable characteristics of recently and acutely infected individuals to guide the prioritization of targeted, differentiated HIV testing, AHI screening, and prevention strategies for these individuals and members of their risk networks. Individuals who are recently or acutely infected can also be offered opportunities to serve as peer mobilizers and/or to distribute HIV self-test kits to help accelerate linkages of their network contacts to testing, treatment, and prevention services, as relevant.

As individuals are screened as potentially having AHI, these individuals can be offered immediate confirmatory HIV RNA testing and HIV

---

**FIGURE 4.** An Illustrative Workflow for Acute HIV Infection Screening, Additionally Applying Recency Testing Data to Help Improve Focus

1. Targeted HIV testing for individuals at high-risk of infection (ideally 4th generation Ag/Ab with verbal acute HIV infection [AHI] screening)
   - Ag & Ab: low verbal screening (Likely uninfected)
   - Ag only or high verbal screening (Suspect acute)
   - Ab+ only (Recent or chronic)

2. Confirmatory HIV testing
   - Case-profiling elicitation
     - Indeterminate (Suspect acute)

3. Support for immediate ART initiation and retention
   - Offer of recency testing algorithm (RITA)
     - Recent infection
     - Chronic infection

4. Index testing and risk network referral
   - High-priority 4th gen index and risk network testing w/AHI screening
   - Index and risk network testing

Abbreviations: Ab+, antibody positive; Ab–, antibody negative; AHI, acute HIV infection; Ag+, antigen positive; Ag–, antigen negative; ART, antiretroviral therapy; HIV+, HIV-positive; PLHIV, people living with HIV.
treatment. On the rare occasion that HIV RNA confirmatory testing is not immediately available, presumptive HIV treatment could be provided pending confirmation in a manner analogous to the provision of HIV postexposure prophylaxis, which is generally considered safe and effective. An advantage of the proposed pooled PCR approach to detect AHI is immediate confirmation of HIV infection. Once confirmation is obtained, individuals can be sustained on ART, having gained personal immunological benefits from early treatment and having reduced the likelihood of ongoing HIV transmission during AHI. For individuals who were screened as having presumptive AHI but who later are determined through HIV RNA testing to be uninfected, treatment can be discontinued with minimal risk of harm or of contributing to development of drug-resistant HIV, in a fashion similar to the discontinuation of HIV postexposure prophylaxis. These persons can also be assessed for the suitability of PrEP.

Current concerns about the potential impact of providing acute or recency test results to clients include increased risk of criminalization of key populations, as well as criminalization of HIV transmission and increased risk of gender-based or intimate partner violence. Furthermore, subjecting patients to tests like recency assays that do not provide additional clinical benefits raises ethical concerns. In principle, patients have a right to know any information that is part of their medical file, and additional information about the current state of a person’s infection may help providers enhance counseling, reinforce a person’s reduction in risk behaviors that lead to onward transmission, improve partner elicitation process within index testing services, and allow providers to use results to prioritize index cases for partner notification services. Adverse events or harm related to return of acute or recency results have not been reported so far from early programs implementing these services, but few studies have systematically evaluated outcomes related to potential harm, client perspectives, or the perspectives of partners of index clients. Given the potential public health benefit of engaging the risk contacts of recently or acutely infected individuals, the assessment of these outcomes is imperative to provide guidance around the messaging of results in a manner which minimizes risks and optimizes potential benefits.

### CONCLUSIONS

Viral burden is the primary predictor of HIV-related morbidity, mortality, and ongoing transmission. Although a majority of PLHIV globally have achieved viral suppression through sustained access to HIV treatment, achieving an end to the HIV pandemic is contingent on addressing the preferences and needs of virally unsuppressed and persistently unserved and underserved PLHIV and members of their risk networks. Proven solutions exist to prevent and treat HIV, but the approaches and technologies needed to differentiate and focus support based on viral burden historically have been limited. Now, with the expansion of viral load testing and an expanded set of options to screen for and treat AHI, we may be better equipped to improve both the impact and efficiency of efforts to accelerate epidemic control.

An HIV micro-epidemic control approach that prioritizes personalized treatment support for PLHIV who are not virally suppressed—and in the process focuses HIV testing and relevant HIV prevention and treatment support among their network members—offers a framework to integrate these advances into current practice to maximize client benefits and overall impact. In particular, such an approach offers a path to integrate the detection and treatment of AHI into routine programming, potentially curbing a substantial proportion of ongoing HIV transmission that occurs during this period and has historically continued apace beyond the reach of efforts to leverage HIV treatment as prevention at scale. However, the ultimate advantages of such an approach remain largely undocumented. Additional investments in the development, implementation, and evaluation of practical strategies to differentiate support based on viral burden are needed to assess the real-world benefit of the proposed HIV micro-epidemic control approach.

**Acknowledgments:** We would like to thank Steve Wignall, Caroline Francis, Rachel Coley, Janet Robinson, Pat Sadate-Ngatchou, Chris Akoko, Meghan DiCarlo, Matthew Avery, and Holly Mahler for their technical contributions to this article.

**Competing interests:** None declared.

### REFERENCES


Capturing Acquired Wisdom, Enabling Healthful Aging, and Building Multinational Partnerships Through Senior Global Health Mentorship

C. Norman Coleman,a John E. Wong,b Eugenia Wendling,a Mary Gospodarowicz,c Donna O’Brien,a Taofeeq Abdallah Ige,d Simeon Chinedu Aruah,d,e David A. Pistenmaa,a,f Ugo Amaldi,g Onyi-Onyinye Balogun,a,h Harmar D. Brereton,a,h Silvia Formenti,a,h Kristen Schroeder,a,i,j Nelson Chao,a,i,j Surbhi Grover,a,k,l Stephen M. Hahn,k James Metz,k Lawrence Roth,a Manjit Dosanjha,m

Key Messages

- Capturing the acquired wisdom and experience of mentors in global health offers a capstone for their careers and provides a purposeful healthspan for these professionals to continue to be engaged in meaningful work while leveraging their expertise to solve challenging health care problems.

- Senior professionals can mentor early career leaders to help them balance their professional commitments, interest in global health, and development of needed skills, such as understanding the nuances of cultural competence and adapting solutions to different environments.

- Institutional leaders, particularly in academic medical centers, recognize the importance of global engagement vis-à-vis their educational mission and for recruiting and retaining faculty and can benefit economically and programmatically from supporting experienced senior faculty or retirees to support these efforts.

- Program builders should include the opportunity for altruistic human service as an integral part of a career and highlight that they can access senior mentors and retirees who provide world-class expertise and mentorship at “volunteer prices.”

INTRODUCTION

An opportunity to have a substantial impact on multiple challenging societal problems exists simultaneously addressing the following: (a) the urgent need for sustainable health care; (b) the importance of mentorship in enabling the emergence of new generations of leaders; (c) the essential need for cross-cultural competency to address global crises through problem solving across societal boundaries; and (d) options for continued productivity by the increasing number of older people. Sustainable health care needs to build on cancer care, which requires urgent intervention and encompasses noncommunicable and infectious diseases in low- and middle-income countries (LMICs) and geographically isolated populations in high-income countries (HICs). Capacity building to meet the cancer care gap, which builds sustainable infrastructure for overall health care and economic development, can be done through twinning programs that engage senior health care professionals in meaningful mentoring roles. As the capstone of a career, these professionals thereby create next-generation leaders within LMICs and their own institutions. This article addresses such opportunities available for individuals in the latter part of their careers including postretirement done either as a continuation of their role as career-long mentors or as a new challenge to be met with their lifelong experience. The expanding and branching tree of mentors to mentees enables a career path in global health and geometric growth to fill in the current enormous capacity gap.

PURPOSEFUL AGING

The challenges facing society regarding the aging of the population are complex. Concepts that have emerged over the past few years to address these challenges include that of “healthspan—the period of life spent in good health, free from the chronic diseases and...
disabilities of aging and the benefit to purpose in life (PIL) for improved health outcomes. Musich et al. noted:

PIL is strongly associated with improved mental and physical health outcomes among older adults. Thus, interventions to improve and/or maintain higher levels of PIL over time may promote successful aging.

This article describes opportunities for professionals to utilize their time and expertise to address the unacceptable gap in cancer care in underserved communities in LMICs and in geographically isolated areas in HICs. Regardless of whether this type of activity promotes longer or healthier lives, it captures expertise that is all too often lost and thereby transfers experience and wisdom to younger generations.

**UNIQUE APPROACH TO THE CHALLENGE OF GLOBAL HEALTH CARE**

Cancer and other noncommunicable diseases (NCDs) represent an increasing share of the global burden of disease in both resource-rich and -poor countries, primarily due to aging, industrialization, sedentary lifestyle, pollution, diet, and the successful approaches to and investment in tackling infectious diseases.

Indeed, addressing the full spectrum of cancer care—prevention, screening, diagnosis, treatment, and long-term follow-up—requires addressing the other major NCDs, such as respiratory, cardiovascular, and metabolic diseases, as well as infectious diseases involved in cancer etiology and those related to treatment. LMICs lack infrastructure, resources, and expertise to address this problem. For example, the workforce shortfall in LMICs is highlighted by the Lancet Oncology Commission’s Global Task Force on Radiotherapy for Cancer Control of the Union for International Cancer Control. Using current staffing models, this report estimates that, by 2035, an additional 30,000 radiation oncologists and over 100,000 technical personnel, as well as clinical support and research staff, will be needed worldwide. The essential health care system expertise and infrastructure needs and the benefits that would be derived from filling these health care gaps make this a formidable and compelling challenge.

Mentorship is recognized as an important element in health care training. Leveraging the expertise and mentorship of senior experts can alleviate this shortfall. Fortunately, because many people are living well past the historical retirement age of 60–70 years, the upward shift of the population age distribution offers a golden opportunity to capture global wisdom, address inequalities, and leverage mentorship and innovative technology to enable sustainable improvement of global health. How best to remain a useful contributor to one’s community and society is a predictable challenge, especially for professionals who have developed the highly sophisticated skill sets required for health care and desire to continue to use their professional knowledge meaningfully. These senior members of a profession also have perspective on the current economic situation in health care, medical and scientific knowledge, and societal trends, as well as broad hands-on patient engagement skills that are particularly relevant in health care in which training and advancement follow a skill-based apprenticeship model.

**PERSON-TO-PERSON CONNECTIVITY AS A SOLUTION SET**

The unprecedented scope of the problems facing humanity today, including climate change, wealth disparities, xenophobia and related terrorism, potential for pandemics, and depletion of natural resources, among others, absolutely requires problem solving across cultures and boundaries. The necessary trusted partnerships/friendships and cultural competence can come from career-long diplomats, altruists, and science-based collaborations, bringing opportunities for groups such as Peace Corps volunteers, professional societies, and nongovernmental organizations. Such organizations span generations, from the eager student to the individuals with decades of experience. The lifelong acquired wisdom of the latter is often lost to retirement, but it is necessary for effective transitions and the transmission of knowledge. Helping those early in their career to visualize a career path in altruistic service can be a powerful motivator and reinforce their own career choices.

A novel approach to address the health care workforce shortfall is the working mentorship model of the International Cancer Expert Corps (ICEC). It draws on a wide breadth of partners and includes the following:

- A collaborative multi-institutional and multinational organization with opportunities for a broad spectrum of experts, who are needed to build an effective health care enterprise to optimize resource utilization and facilitate the transfer of professional and technologic experience and expertise
- Assignments in established and emerging twinning partnerships with HIC expert academic
centers, professional societies, and private practices mentoring programs in LMICs, thereby offering long-term guided progress as opposed to episodic visits.

- Tools and resources to guide mentoring and program-building efforts including standard operating procedures, the detailed metrics in the ICEC 5-Step Progression Plan for Cancer Care, and formal guidelines for education and training programs for global settings.

- Ways to contribute expertise to support volunteer education programs such as Chartrounds’ case conferences for LMIC participants.

- Opportunities for mentors to get formal recognition for their contributions, as part of a shared mission, while assisting in the development of a career path in global health.

Expertise can come from both people and technology. For health care in developing countries and for developed countries in the future, where rising expenses are a major societal issue, building human and technology expertise together, using the rapidly growing area of artificial intelligence and machine learning, can better utilize human resources. Technology requires appropriate training and support services. The teams providing care in some poorly resourced countries may have access to excellent (highly publicized and often very expensive) equipment, but they may not have the expertise to fully utilize it. This problem is being addressed by ICEC and its LMIC partners, Medical Physics for World Benefit, the International Atomic Energy Agency Division of Human Health, and academia.

The Figure illustrates the mentorship model for patient-centered cancer care, which encompasses a broad range of expertise including NCDs and infectious diseases. Mentorship includes the continuum of mentors, with senior mentors guiding early- and mid-career mentors from well-resourced programs (hubs) who jointly train and educate mentees and staff within LMICs and geographically isolated regions in HICs (centers), thereby geometrically expanding the system of patient-centered care. Senior expertise, a very expensive component of health care (“Solution shop” of Christensen et al.), can be made available much less expensively with this sustainable volunteer mentorship approach. Sharing knowledge and broad expertise in this manner enhances and expands their value well beyond the one-to-one mentor-mentee relationship. This innovative paradigm captures acquired wisdom, which is often lost following retirement, to benefit society.

### TWINNING: MENTOR-MENTEE PARTNERSHIPS

The mentorship model illustrated in the Figure works primarily through twinning programs that...
are collaborative relationships between HIC university departments or private practice programs (hubs) and programs/facilities in an LMIC (centers). The value of mentorship and the exponential impact of transferring experience are apparent in the twinning programs that establish a proper infrastructure for education, training, and mentoring. This capacity-building strategy facilitates the creation of a sustainable platform for the mutual sharing of best practices and learning through information and technology transfer. The ultimate aim is for the centers to achieve the required level of expertise to become hubs for their respective regions. A successful international pioneering example is the King Hussein Cancer Center in Jordan, which is now a regional leader in cancer care. Mentoring at the trainee level is exemplified by the work of the Association of Residents in Radiation Oncology Global Health Initiative. For problems as large as the gap in global health care that may seem "too hard" to address, specific examples can make the solution less daunting and even an exciting personal challenge.

Because going from concept to operational reality is critical, we include a narrative example of successful mentorship from a mentor and a mentee (Box) (additional examples are included in the Supplement). An important starting point is that even pursuing a sustainable career in global oncology had been a challenge, yet these mentor-mentee teams have opened up this possibility to an emerging generation committed to global health. The mentor-mentee model has already demonstrated success, as shown with examples in Table 1.

Dr. Onyinye Balogun, a radiation oncologist from Weill Cornell Medical School, has established training programs in Armenia that have enabled the radiation oncologists to jump forward a few decades in radiation oncology from 2-dimensional radiation therapy to 3-dimensional techniques. Training and ongoing telemedicine case discussions enable further advancement in techniques that are less toxic and, by allowing higher doses, more effective. Her work and that of her mentors led the dean to establish a global oncology initiative at the medical school.

Surbhi Grover, MD, completed an MPH degree under mentorship advice and with support from the University of Pennsylvania. She has been hands-on in Botswana establishing evidence-based cancer care guidelines. This work is a major advance in care and has transformed the strategies to manage stock for chemotherapy as part of comprehensive care plans. As one of the first radiation oncologists to be on the ground in global health, her program is a highly sought-after rotation for residents interested in pursuing careers in global health.

Kristin Schroeder is a pediatric hematologist-oncologist from Duke University who has helped establish pediatric cancer care in Tanzania. She has not only implemented a comprehensive cancer care infrastructure, but also helped establish a nongovernmental organization to provide care to any child with cancer.

Taofeeq Abdallah is a medical physicist in Nigeria who has established, under the mentorship of CERN, education and training networking...
Mentee: Taofeeq Abdallah, IGE, PhD, Chief Consultant Physicist and Head of Medical Physics Department, National Hospital, Abuja, Nigeria

Mentor: Manjit Dosanjh, PhD, Senior Advisor for Medical Applications, European Organization for Nuclear Research (CERN), Geneva, Switzerland

Taofeeq Abdallah: The relationship between Manjit and myself since our first meeting in CERN in 2017 has been more than awesome. She brought a fresh perspectives between a mentor and mentee by trying to identify real-time with the situation in the LMIC’s and this has propelled me and my colleagues on this “side of the divide” to push ahead even more vociferously knowing fully well that we can always rally for support anytime that this is needed and she has never disappointed in all the occasions – always rising up to the challenge and offering advice that are most accurate and incisive.

The tangible benefits that this international mentoring relationship have engendered has been first to our numerous patients who have in one way or the other benefited from very rich advice that Manjit has been able to offer from time to time – raising our spirits even in the face of arduous and unfavorable conditions. Since the relationship impacts our patients, this has equally been of great benefit to me professionally and has had a concomitant net benefits to my hospital and even my interactions with colleagues in the region as the president of our professional association (FAMPO – Federation of African Medical Physics Organizations).

Manjit Dosanjh: I got to work with and to know Taofeeq much more closely when the Science and Technology Facilities Council (STFC) team started to prepare a proposal to conduct an Accelerator Design Study (ADS) for a medical linear accelerator (LINAC) for Overseas Development Agency countries to be submitted to the Global Challenges Research Fund.

At my suggestion, both Taofeeq and Simeon Chinedu Aruah were invited to participate in the preparation of the ADS to advise the STFC team about both clinical and medical physics challenges of LINAC use in Nigeria. During the period of the development of the ADS proposal, I realized that Taofeeq and Simeon were not used to communicating and working closely with each other. This fact provided a great opportunity for me to help bridge that gap and build a closer working relationship between them.

Since then, I have been guiding Taofeeq in how to prepare and submit his own projects; he led the last one with myself as a co-applicant. We are now working on a questionnaire gathering information for optimizing a LINAC prototype for future machines suitable for challenging environments. Also, David Pistenmaa and I accepted Simeon and Taofeeq’s invitation to contribute to peer-reviewed manuscripts that they originated and enjoyed the camaraderie in doing so. What has been most rewarding to us over the last 2 years has been to see not only Taofeeq become a more understanding and caring leader but also to see the relationships between him and Simeon and their departments growing. These improving collaborations will continue to enhance the quality of treatment of patients with cancer and the reputation of National Hospital Abuja.
<table>
<thead>
<tr>
<th>Mentee</th>
<th>Mentor</th>
<th>Year Started</th>
<th>Examples of Achievement</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Center of Oncology, Yerevan, Armenia, and Weill Cornell, New York, USA</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Onyinye Balogun, MD | Silvia Formenti, MD | 2015 | • Established a training program to facilitate transitioning from 2-dimensional to 3-dimensional treatment planning for treatment of cancer with radiotherapy, with a focus on breast cancer  
• Established education and ongoing training program to ensure proper implementation of image-guided brachytherapy for cervical cancer. Training is delivered through didactic lectures and teleconferences offering patient case discussion and peer review  
• Established the global oncology initiative at Weill Cornell Medicine  
• Established one of the first ICEC twinning programs linking an emerging cancer treatment program in an LMIC with an advanced cancer treatment program in an HIC |
| Assistant Professor of Radiation Oncology, Weill Cornell | Chairman Department of Radiation Oncology, Weill Cornell  
Associate Director of the Meyer Cancer Center and Radiation Oncologist in Chief, New York Presbyterian Hospital | | |
| Harmar Brereton, MD | Clinical Professor of Medicine  
Geisinger Commonwealth School of Medicine and Clinical Assistant Professor of Radiation Oncology, Weill Cornell | | |
| Princess Marina Hospital, Gaborone, Botswana, and University of Pennsylvania, Philadelphia, Pennsylvania, USA | | | |
| Surbhi Grover, MD, MPH | Stephen Hahn, MD | 2014 | • Increased evidence-based care establishing guidelines for the top 10 cancers in Botswana  
• Created an educational exchange program between University of Botswana and University of Pennsylvania  
• Developed research programs between the University of Botswana and University of Pennsylvania Radiation Oncology expanding research capacity at University of Botswana and linking young investigators to international mentors to support research  
• Advanced strategies to reduce stock-outs of chemotherapy and to improve systems to reduce delays in pathology diagnosed through an initiative with the American Society of Clinical Pathology  
• Botswana is now a destination for radiation oncology residents pursuing careers in global health, orchestrated by the Association of Residents in Radiation Oncology |
| Assistant Professor of Radiation Oncology, Perelman School of Medicine, University of Pennsylvania  
University of Botswana & Princess Marina Hospital, Gaborone, Botswana | FDA Commissioner  
Former Chair, Department of Radiation Oncology Perelman School of Medicine, University of Pennsylvania | | |
| James Metz, MD | Chair, Department of Radiation Oncology, Perelman School of Medicine, University of Pennsylvania | | |

Continued
<table>
<thead>
<tr>
<th>Mentee</th>
<th>Mentor</th>
<th>Year Started</th>
<th>Examples of Achievement</th>
</tr>
</thead>
</table>
| Bugando Cancer Center, Mwanza, Tanzania, and Duke Children’s Hospital & Health Center, Durham, NC, USA | Kristin Schroeder, MD, MPH  
Pediatric Hematology-Oncology Specialist, Pediatric Neuro-oncologist  
Bugando Cancer Center, Mwanza, Tanzania  
Duke Children’s Hospital & Health Center, North Carolina, USA | 2014          | - Established a patient navigator program that offers education and caregiver guidance throughout the diagnosis and patient treatment  
- Developed a pediatric cancer clinical database to monitor patient outcomes  
- Established a hospital-based cancer registry  
- Fostered a streamlined process to speed cancer diagnosis and access to treatment  
- Implemented standard protocols for care  
- Initiated research programs related to Burkitt lymphoma and retinoblastoma treatment, and impact of psychosocial support  
- Founded the NGO, International Cancer Care and Research Excellence Foundation (iCCARE), a nonprofit whose mission is to give any child diagnosed with cancer the same chance of a cure regardless of where they live  
- Her mentorship of 19 individuals includes 2 Fulbright scholars, 5 masters level students, 1 oncology fellow, 2 nurses, 1 resident, 3 medical students, and 3 undergraduate students |
| National Hospital, Abuja, Nigeria, and the European Organization for Nuclear Research (CERN) | Taofeek Abdallah Ige, PhD  
Chief Consultant Physicist and Head of Medical Physics Department, National Hospital, Abuja, Nigeria | 2017         | - Established cross-border networking, education, and research projects to enhance the accessibility, effectiveness, and safety related to the use of medical physics and technologies improving treatment techniques and patient outcomes  
- Fostered mentoring relationships between individuals in HICs and LMICs providing access to expert knowledge, guidance, advice, and building collegial relationships  
- Established knowledge- and information-sharing programs utilizing various platforms including WebEx and videoconferencing and attendance at global scientific meetings  
- Facilitated engagement in research programs resulting in co-authorship on scholarly articles published in leading academic journal publications |
TABLE 1. Continued

<table>
<thead>
<tr>
<th>Mentee</th>
<th>Mentor</th>
<th>Year Started</th>
<th>Examples of Achievement</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Hospital, Abuja, Nigeria, and the International Cancer Expert Corps (ICEC)</td>
<td></td>
<td></td>
<td>• Fostered academic growth through co-authorship on publications in top scientific oncology journals&lt;br&gt; • Presentation of quality papers in different fora&lt;br&gt; • Capitalized on opportunities to travel outside Nigeria to attend international workshops, which has widened access to world class education and training, resulting in improved delivery of quality cancer care in Nigeria&lt;br&gt; • Increased global visibility of National Hospital Abuja through representing Nigeria in the 63rd International Atomic Energy Agency general assembly in September 2019 in Vienna, Austria, and an invitation to represent Nigeria at the UN Disarmament Conference in New York City in May 2020 (postponed because of COVID-19)&lt;br&gt; • Increased respect and enhanced image of the National Hospital Abuja within the scientific community&lt;br&gt; • Improved the quality of academic lectures to resident doctors and undergraduate medical students resulting in the fostering of new mentoring relationships within Nigerian hospitals and academic medical centers</td>
</tr>
<tr>
<td>Simeon Chinedu Aruah MD, MPH, FWACS&lt;br&gt;Consultant Radiation and Clinical Oncology&lt;br&gt;National Hospital Abuja, Nigeria&lt;br&gt;Lecturer University of Abuja College of Medicine, Nigeria&lt;br&gt;Head of Department Radiation and Clinical Oncology&lt;br&gt;National Hospital Abuja, Nigeria</td>
<td>David A. Pistenmaa, MD, PHD, FACR&lt;br&gt;Chief Scientific Program Director&lt;br&gt;International Cancer Expert Corps&lt;br&gt;Manjit Dosanjh, PhD&lt;br&gt;Senior Advisor for Medical Applications, CERN</td>
<td>2017</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: COVID-19, coronavirus disease 2019; HIC, high-income country; ICEC, International Cancer Expert Corps; LMICs, low- and middle-income countries; NGO, nongovernmental organization.
to enhance safety and teach a sophisticated technique for radiation therapy.

Simeon Aruah, MD, MPH, is a young lecturer in radiation and clinical oncology (includes medical oncology) in Nigeria, who has had a rapid growth in his academic career with assistance in conducting studies and preparing manuscripts and presentations from his mentors. His linkage with world-renowned academic mentors increased the visibility of his program and cancer care in Nigeria. His talent, enthusiasm, and confidence have grown, and he has already represented Nigeria at the International Atomic Energy Agency general assembly and will do so at the UN Disarmament Conference in New York City.

Interest in careers involving global oncology has surged with these pioneering examples, including program leaders willing to support trainees and faculty as part of a career path. The experience of the co-authors of this article can attest to the positive impact that the mentor-mentee relationship has on stimulating transgenerational idea sharing and generating energy and a positive outlook for what can be done, despite challenges that appear discouraging.

**OPPORTUNITY FOR A BROAD RANGE OF EXPERT MENTORS**

Improved health care in general, building on the spectrum of cancer care from prevention through diagnosis, treatment, and long-term follow-up care, are the deliverables. Enabling this goal requires contributions including and well beyond patient care delivery from a broad range of experts, as shown in Table 2.

Critically, the focus on cancer as part of the health care system encompasses the other NCDs—respiratory, cardiovascular, and metabolic diseases—and infectious diseases that are linked to both cancer etiology and complications of treatment. Cancer is a logical focus in that it has a sense of urgency, similar to infectious diseases, and it can be a focal point for community involvement. Thus, an opportunity exists for volunteers with a wide range of skills and expertise, including medicine, a broad range of scientific disciplines, and other professions (“broad support” such as cultural experts, communications, logistics, finance, and legal), to effectively transfer knowledge and wisdom, while reducing the expense associated with personnel. Thus, global mentorship teams can educate one another and provide mentorship to the local champions who are building programs in underserved communities, enabling the geometric expansion of health care necessary to address the enlarging workforce shortfall. Cultural competency is essential, and it benefits from those with in-country experience. Answering the questions of “What can I do, and how do I do it?” is facilitated by structure with achievable expectations. For a mentor, the expectation is only a ~20% time commitment (8 hours per week on average), with the vast majority of the mentoring by planned teleconferencing (with some bidirectional travel possible) through protocol- and guideline-based care, rather than individual case management. As described by Crisp, knowledge and models for care will also evolve from mentees to mentors through reverse innovation.
In our opinion, raising the mandatory retirement age or eliminating it altogether—as is happening in many societies—presents new opportunities for those affected by necessary transitions in leadership in health care organizations, governmental and international diplomatic organizations, and academia that free senior personnel to mentor within step-down roles at work or later in retirement. This transition of lives and careers provides exceptional opportunities for the older generation to pass on its knowledge and wisdom, while enhancing their own quality of life. This sentiment was recently highlighted by Jane E. Brody in "Want to leave a legacy? Be a mentor." For those interested in a continued purpose in life related to their profession, addressing both healthspan and lifespan requires opportunities to use their skills by volunteering time and expertise. This trend ("purposeful healthspan") utilizes this experience at a low cost for a wide range of organizations and interest groups, such as professional societies, religious organizations, and specific social causes. The model presented here is based upon periodic short visits followed by sustainable commitments and continuous mentoring of those working on site through teleconferencing. The model creates long-standing relationships and programs built around the broad range of skills transferred to the mentored specialists and staff in the local communities (Table 2).

Several similar large-scale mentorship model programs have found the keys for successful implementation to be convenience, flexibility, and purposefulness. Such examples include the following: (1) The AARP Foundation Experience Corps’ intergenerational volunteer-based tutoring program designed to help elementary school students improve their reading levels and to help older adults enrich their lives through literacy; (2) the Returned Peace Corps Volunteers; (3) the Japanese government’s “The Community-based Integrated Care System” providing comprehensive up-to-the-end-of-life support in every community; (4) Singapore’s Action Plan for Successful Aging that enables seniors to learn new skills in joyful endeavors and to deploy these skills, and (5) the National University Health System and the National University of Singapore’s work with multiple government agencies to enable whole precincts to exploit the elements of successful aging, thus future-proofing Singapore as a livable city for people of all ages. This “whole of society” and “whole of government” approach will enable societal change to take place and remain sustainable.
CONCLUSION

The confluence of opportunities for continuity and the spectrum of expertise from senior mentors to those early in their career has not been more apparent than in the current COVID-19 pandemic. On the one hand, the call for retirees to return to health care speaks to how senior and world-renowned experts’ skills are useful for their primary expertise and for their role modeling, gravitas, and potential direct support. Yet, on the other hand, greater awareness of such usefulness and the presence of senior experience and wisdom might have averted the wholesale dismissal of 7,300 Peace Corps volunteers and fostered a more appropriate transition.

Older individuals have opportunities to serve society and humanity. Such opportunities (1) provide a career capstone, (2) allow timely transfer of institutional responsibility to next-generation leaders, (3) establish mentorship relationships for world-renowned experts with dedicated professionals in underserved and geographically remote health care regions, (4) provide expensive expertise at “volunteer prices,”11,20 (5) present a model for geometric expansion of diverse expertise and innovative technology that enables development of the capacity to effectively address the burgeoning burden of cancer and other NCDs, (6) establish a mentor-based career path for altruistic human service that is an endangered species in the current “bottom line” finance-driven health care system, (7) emphasize the importance of cultural competence and listening, and (8) utilize a systems solution approach to improve health care in LMICs by developing and sustaining local champions. The presence of a gap that can be filled in a rather short timeline from mentor to mentee to LMIC mentee speaks to the need and impact.

Whether being engaged in purposeful activities, such as those described in this article, will increase the length of one’s lifespan is under study. Such study includes understanding the impact of aging on the workplace.36 Interestingly, coinciding with this current article, Dzau et al.51 recently announced “The National Academy of Medicine Grand Challenge in Healthy Longevity.” What is undeniable is that the benefit of such activities can meaningfully increase the breadth of one’s experiences and contributions to society. Serving as a senior mentor to mentees in resource-poor regions of the world can have a spectacular impact on the goal of rectifying the staggering lack of access to care for patients with cancer and other NCDs in those regions. In addition, by improving the infrastructure of cancer care facilities and expanding the breadth of expertise available to them, these facilities can serve as focal points for the development of sustainable on-the-ground programs that can have substantial health and economic benefits beyond cancer care. Transformational models, as outlined in this article, offer opportunities for visionary investments, altruistic contributions, and exciting and meaningful action for a purposeful aging and improved healthspan.

Competing interests: None declared.

REFERENCES


Peer Reviewed

Received: March 7, 2020; Accepted: August 24, 2020; First published online: October 15, 2020


© Coleman 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: http://doi.org/10.9745/GHSP-D-20-00108

Kasereka M. Claude, a Muyisa Sahika Serge, a Kahindo Kahatane Alexis, b Michael T. Hawkes c–g

Key Findings

- Congolese internally displaced persons (IDPs) had high awareness and fear of COVID-19, but low specific knowledge.
- IDPs face major barriers to implementing COVID-19 prevention measures: crowded shelters, frequent movements in and out of the camp for work, and lack of soap for hand hygiene.
- IDPs’ desire for peace and to return to their native homes, where COVID-19 precautions could be feasibly implemented, overshadowed their enthusiasm for other control measures such as a vaccine.

Key Implications

- Donors and policy makers should consider providing consumables, such as soap for hand hygiene and face masks, to implement COVID-19 precautions.
- The national government or international aid agencies should consider providing individual family dwellings (e.g., tarpaulin tents) to allow IDPs to practice physical distancing.
- National and international governments should take serious measures to restore peace to the area by controlling armed conflict. A safe return to their homes would allow IDPs to practice COVID-19 prevention without external aid.

Abstract

Background: The coronavirus disease (COVID-19) pandemic poses a grave threat to refugees and internally displaced persons (IDPs). We examined knowledge, attitudes, and practices with respect to COVID-19 prevention among IDPs in war-torn Eastern Democratic Republic of the Congo (DRC).

Methods: Mixed-methods study with qualitative (focus group discussions, FGDs) and quantitative (52-item survey questionnaire) data collection and synthesis.

Results: FGDs (N=23) and survey questionnaires (N=164 IDPs; N=143 comparison group) were conducted in May 2020. FGD participants provided narratives of violence that they had fled. IDPs were statistically more likely to have larger household size, experience more extreme poverty, have lower educational attainment, and have less access to information through media and internet versus the comparison group (P<.05 for the comparison group). IDPs had a high level of awareness (99%) and fear (98%) of COVID-19, but lower specific knowledge (15% sufficient knowledge versus 30% among the comparison group, P<.0001), a difference which remained significant in a multivariable model adjusting for confounding. IDPs faced major barriers to implementing COVID-19 prevention measures. Physical distancing was impossible for IDPs in crowded shelters, and 70% reported coming in close contact with someone other than a family member within the past 24 hours (versus 56% of the comparison group, P=.014). Frequent movements in and out of the camp for subsistence left IDPs vulnerable to the introduction of COVID-19: 61% left the camp on a daily basis and 65% had received a visitor in the past month. Despite acceptance of hand hygiene for prevention, 92% lacked soap (versus 65% of the comparison group, P<.0001). IDPs’ desire for peace and to return to their native homes, where COVID-19 precautions could be feasibly implemented, overshadowed their perceived benefits of measures such as a COVID-19 vaccine.

Conclusions: These findings provide empiric evidence supporting the vulnerability of IDPs to COVID-19 and call for action to protect neglected displaced populations.
the United States and Europe experiencing the highest burden.\textsuperscript{1} African countries have reported 298,000 cases and 8,000 deaths (case fatality ratio 2.4%).\textsuperscript{2} Many low-resource settings lack comprehensive surveillance and laboratory testing to monitor the spread of COVID-19.\textsuperscript{3} The presence of displaced populations (refugees and internally displaced persons [IDPs]) adds further complexity to the COVID-19 pandemic and control measures in low- and middle-income countries (LMICs) in conflict zones.

In the Democratic Republic of the Congo (DRC), the first case of COVID-19 was detected on March 10, 2020, in a traveler returning from France.\textsuperscript{4} Since then, more than 9,800 cases and 251 deaths have been confirmed across the DRC. Most cases have been detected in the capital city, Kinshasa. In the province of North Kivu, there have been 203 cases as of August 25, 2020. The primary mode of transmission is community based.\textsuperscript{5} In response to the pandemic, the government declared a state of public health emergency on March 24, 2020, with broad closure of businesses, gatherings, and travel.\textsuperscript{4} Since this initial lockdown, the government authorized gradual reopening of businesses and public transportation (July 22); schools and universities (August 3); and churches, interprovincial travel, and international airports (August 15).\textsuperscript{4}

Refugees and migrants are among the world’s most vulnerable people.\textsuperscript{6} Worldwide, there are approximately 26 million refugees and 46 million IDPs, displaced due to insecurity and natural disasters.\textsuperscript{7} The DRC has the second highest number of IDPs of any country in the world (after Syria), estimated at more than 5.5 million.\textsuperscript{8} Displaced populations, housed in temporary shelters or camps, generally have limited access to quality shelter, sanitation, clean water, stable food supply, and health care. Under these conditions, COVID-19 prevention efforts may be challenging.\textsuperscript{9-11}

Impacts of the COVID-19 pandemic on displaced populations are predicted to be disastrous. Already, resettlement procedures have been suspended by the United Nations, alongside widespread travel bans. The first case of COVID-19 in the island of Lesvos in March 2020 raised the alarm for the 20,000 residents of the Moria refugee camp, where distancing is a physical impossibility.\textsuperscript{9} In the world’s largest refugee camp in Bangladesh, which shelters more than 855,000 Rohingya refugees, preparations for COVID-19 have begun, such as portable handwashing facilities at every community center.\textsuperscript{12} In Nigeria, efforts to mitigate the impacts of COVID-19 have included sensitization campaigns on handwashing and distribution of soap to more than 100,000 IDPs.\textsuperscript{13} However, as noted by previous authors, recommendations for hand hygiene and physical distancing may be extremely difficult to implement in a refugee or IDP camp. How do you self-isolate in a refugee camp?\textsuperscript{10} Several commentators have forewarned of an impending crisis if COVID-19 strikes in refugee or IDP camps.\textsuperscript{6,10,12} However, a paucity of empirical data from these areas is available.

Our overarching goal was to contribute to the improvement of prevention strategies against COVID-19 in IDP camps in the DRC. We aimed to describe the knowledge, attitudes, and practices (KAPs) of IDPs in Eastern DRC with respect to the prevention of COVID-19. Our primary endpoint was COVID-19 specific knowledge, which we compared between IDPs and individuals from neighboring villages. Other specific objectives included: (1) to describe attitudes of IDPs with respect to COVID-19 and its prevention; (2) to describe the practices used by IDPs for preventing COVID-19; and (3) to describe barriers faced by IDPs in implementing recommended COVID-19 prevention measures.

### METHODS

#### Study Design

We conducted a mixed-methods study with qualitative focus group discussion (FGDs) and quantitative (52-item survey questionnaire) data collection. Mixed-methods research seeks to triangulate data from qualitative and quantitative methods.\textsuperscript{14} Convergence of findings from multiple methods may enhance the validity of results (multiple operationalism).\textsuperscript{15} We and others have previously used this methodology to integrate community attitudes, behaviors, and responses into epidemiological research.\textsuperscript{16,17} With respect to the survey questionnaire, the study followed a descriptive cross-sectional design.

#### Study Setting

The province of North Kivu has a population of 6.7 million inhabitants and an estimated 1.7 million IDPs.\textsuperscript{18} The Eastern provinces of the DRC have been the arena of a complex humanitarian emergency for several decades. Mortality rates are 70% above pre-war levels, due largely to preventable and treatable infectious diseases rather than the direct effects of conflict.\textsuperscript{19} Large-scale population displacement has resulted in numerous IDP...
camps throughout the area. The chronic threats to security have long been neglected by the national government and the international community.

We selected 3 IDP camps (Mwangaza, Masosi, and Luvangira) located 2 to 5 km from the rural commune of Oicha, North Kivu. These temporary settlements consisted of groups of IDPs sheltered in school buildings or mud/thatch dwellings on public grounds. Camp census data indicated the following populations: Mwangaza (1064 individuals, 200 households); Masosi (869 individuals, 176 households); and Luvangira (250 individuals, 75 households). Aid for the camp is coordinated by the nongovernmental organization Charité Aide et Développement, Axe Oicha, with intermittent assistance from OXFAM, World Food Programme, and International Committee of the Red Cross.

**FGDs**

Participants of FGDs were purposively selected from the 3 IDP camps. Participants included adult women (3 FGDs) and men (2 FGDs) who were heads of households, and youth (1 FGD). FGDs were conducted in Congolese Swahili. Discussions were recorded, translated, and transcribed into English for subsequent analysis. FGDs lasted 30–45 minutes and included 3 or 4 participants in each group. The FGD topic guide was adaptive, allowing us to confirm findings and explore emerging themes from each FGD session. Questions were open-ended and elastic, allowing participants to shape the discussion. FGDs were continued until saturation. Thematic analysis was used to identify, analyze, and report themes in the FGDs. Two investigators (KMC and MH) read the transcripts several times, noted preliminary ideas, produced initial codes, then generated and refined themes. Representative quotations as well as statements of particular interest were extracted to support the themes.

**Survey Questionnaire**

We developed a 52-item questionnaire based on past COVID-19 questionnaires used in Guyana and Uganda. The choice of questionnaire items was guided by a need for contextually appropriate questions for low-income settings. We also drew on past experience from past surveys conducted in IDP camps in the area and from the recent Ebola virus disease epidemic to design questions that would be relevant and understood by the participants. A local Congolese physician (KMC) with tacit knowledge of the circumstances, culture, and language of the IDPs chose the appropriate wording of the questions and adapted the content of the questionnaire to the conditions in the IDP camp. The survey was administered to IDPs in the 3 camps as follows.

**Sampling**

**Statistical Unit and Estimation of Sample Size**

The unit of analysis was the household, defined as a family unit, often consisting of male and female parents and their children.

For our primary analysis, we focused on differences in sufficient knowledge (binary variable) between IDPs and the comparison group. A standard sample size calculation indicated that 138 households would be needed to detect a difference of 15%, with 95% confidence and 80% power, assuming that the proportion of IDPs with sufficient knowledge was 20% or less, based on our previous study of knowledge of Ebola virus disease among IDPs.

**Sampling Technique**

Geospatial sampling was used, as in previous studies of mobile populations. IDP camps were divided geographically into thirds and 1 area was chosen at random. All households living within the selected area were included, and the standardized questionnaire was administered to 1 adult member from each household. Our sampling technique was inspired by the cluster sampling method developed by the WHO for monitoring vaccine coverage. In this approach, a population is divided into a specified number of geographic “clusters” (in our case, camps) of a known or estimable population size. Within each cluster, the desired number of households are selected (in our case, approximately one-third were needed to reach the required sample size). Several strategies are possible for household selection (e.g., enumeration of all households and simple random sampling from this list, or a “random walk” sampling contiguous households). However, random selection in more densely populated areas (e.g., urban settings or, in our case, an IDP camp) can be more challenging, given the more complex household types (e.g., apartment buildings or, in our case, IDPs sheltering in school classrooms). In such settings, a common approach is to divide the geographic area of interest into zones, randomly select a zone, and randomly select a starting point within that zone. To reach our desired sample size, we needed to sample.
approximately one-third of the camp households. Therefore, we chose to divide the camp into thirds, choose 1 cluster at random, and sample all households within that cluster. For our comparison group, we surveyed the surrounding villages (nondisplaced population) using a nonprobability, purposive, maximum variation sampling technique, choosing participants from all demographic categories (men and women, full age spectrum, employment category, education attainment, and marital status). Participants were aged 18 years or older.

List of Variables
The questionnaire consisted of several domains related to participant demographics, knowledge of COVID-19, attitudes, and behaviors for preventing COVID-19.

Demographics
Individual respondent characteristics were collected: age, sex, educational attainment, and marital status. In addition, we collected data on household characteristics (number of family members, members aged 60 years and older) and wealth indicators (ownership of radio, cellular telephone, and bicycle).

Knowledge of COVID-19 Symptoms
Participants were asked to choose from a list of possible sources they drew upon for information on COVID-19 (multiple selections possible). Using a list of symptoms, including 2 detractor (false) symptoms (constipation and bleeding), participants were asked to agree whether COVID-19 was associated with each symptom (“yes” or “no”). Recognition of asymptomatic transmission38,39 was assessed with the question: “A person who is not sick and who has no symptoms can still spread the virus” (responses: “true,” “false,” or “I don’t know”). Agreement with common misconceptions (transmission by mosquitos, prevention with spicy food) was assessed (responses: “yes,” “no,” or “I don’t know”). Participants were considered to have sufficient knowledge of COVID-19 if they identified at least 1 of the cardinal signs and symptoms of COVID-19 (fever, cough, or difficulty breathing).40

Attitudes
We probed a range of attitudes related to COVID-19 by assessing agreement with statements on a 5-point Likert scale (“strongly agree,” “somewhat agree,” “neutral,” “somewhat disagree,” and “disagree.” with a possible “I don’t know” response). Affective response was measured using 2 questions about perceived severity and fear of COVID-19. We assessed attitudes toward recommended control measures, including physical distancing and staying home without working. Mistrust and rumors contributed to community resistance to control measures during the recent Ebolavirus epidemic in the DRC.17,41 Therefore, we included measures of institutional trust (2 items) and endorsement of conspiracies related to the SARS-CoV-2 virus (2 items).

Data Collection Technique
The standardized questionnaire was administered as a verbal structured interview, with a study team member asking questions in the local language and recording the participant’s answers using a field-adapted electronic data collection tool, KoboToolbox.42 Study team members were local Congolese health workers with tacit understanding of the language and culture, biomedical understanding of COVID-19, and past experience administering surveys by verbal interview.

Data Processing and Analysis
For descriptive statistics, we used median and interquartile range for continuous variables, and number (percentage) for proportions. Comparative statistics were computed using non-parametric methods: Mann-Whitney U-test for continuous variables and Chi squared or Fisher’s exact test for dichotomous variables, as appropriate. With respect to our primary analysis, we expressed the
association between IDP status and knowledge as odds ratio (OR), the cross-product ratio of the entries in the 2-by-2 contingency table of 2 binary variables. Multivariable logistic regression was used to verify the association between IDP status and knowledge, with adjustment for confounding variables. Statistical analysis was performed in the R statistical environment.

Ethics Considerations
Participants provided verbal consent to participate in the FGD and the questionnaire. Ethics approval was obtained from the Comité d’Éthique du Nord Kivu (Université Catholique du Graben, ref 003/TEN/CENK/2020). Operational approval was granted by the municipal authority (bourgmestre) and the local refugee coordinator. Participant confidentiality was respected during implementation and analysis of survey results. Data were collected anonymously, without identifiers, and all results were presented in aggregate so that no individual participant can be identified. All names and locations were removed from FGD quotations to avoid possible identification of the speaker.

RESULTS

FGD Themes
We began with a qualitative exploration of COVID-19 prevention in the IDP camp. We conducted 6 FGDs, involving 23 participants (total). The composition of focus groups is shown in Table 1. FGDs generated rich qualitative data, from which we derived the following themes: (1) displacement narratives; (2) population movements in and out of the camp and risk of introducing COVID-19; (3) high level of awareness and fear of COVID-19; (4) challenges associated with hand hygiene in the camp; (5) impossibility of physical distancing in the IDP camp; and (6) restoring peace and security takes priority over vaccine.

We elaborate on each theme and provide representative quotations.

Displacement Narratives
Unspeakable terror and killings drove FGD participants from their native homes.

I’ve been in this camp for 6 years, since the beginning of the massacres in the region. —FGD2, F1

We called them the “ba chinja chinja” [throat-slitters]. —FGD1, M2

We saw serious atrocities and these will stay in our memories for a long time. —FGD2, F1

The insecurity has now become permanently established there. They killed people there, including members of my family. —FGD1, M2

Me, I don’t like to be reminded of this. We suffer a lot. —FGD3, M2

The journey IDPs had followed to reach the camp was often challenging and circuitous, passing through multiple temporary dwelling places before arriving in their current camp:

We spent nights outside in the bush during the armed attacks by those people. —FGD5, F7

In reality, when these people come to kill, you are just driven by a reflex to survive initially. And the next day, you ask yourself: now what? What do I do? —FGD1, M1

First, it’s panic, you have to flee and you don’t know who is where. You leave the house empty-handed, maybe with a child, and everybody has to flee. The next day, it’s counting the dead and the damages. Then rapidly finding where to stay for security. —FGD1, M3

We passed through several areas, depending on the security situation. There was a lot of back and forth just attempting to restart a stable life. —FGD1, M4

Loss of housing, assets, and livelihood meant that IDPs current condition was precarious:

The war… a very bad thing. They attacked my village several times and we had to abandon everything, eventually arriving here. —FGD3, M5

Those fields are our guarantee for life. —FGD1, M1

...Our saving for the present and the future. It’s our wealth, what keeps us alive, feeds us, pays for health care and school for our children. —FGD1, M4

Some FGD participants expressed paralysis, hopelessness, and a sense of abandonment:

On 1 side, the insecurity, and on the other, this corona—yes, we are scared. I’m just in shock. I can’t say anything at the moment. And tell that government, there, that we are abandoned here. —FGD5, F5

Population Movements In and Out of the Camp/ Risk of Introducing COVID-19
Some FGD participants pointed out the insecurity and isolation of the camp that restricted travel:
The area is very dangerous, they try to limit the movements. —FGD4, Y3

It’s rare to visit others. We spend most of the time here or in the fields. —FGD4, Y3

Others identified sources of visitors from outside the camp and noted that many IDPs move out of the camp for work on a daily basis.

They [visitors] come from other camps or people who have fields that employ us to work in their fields. —FGD2, F2

No, in terms of leaving the camp, you can’t count the number of times. If you stay here, the children will die of hunger. Many times a day to look for something to eat. To the market, to the fields, anywhere that you can find something. —FGD6, F10

There is a constant coming and going of people from outside the camp and vice-versa. —FGD 6, F11

High Level of COVID-19 Awareness and Fear

There was a high level of awareness and fear of COVID-19, which was known as “corona”:

A new disease and very severe. We are afraid of it and we pray that it stays away from us. —FGD3, M7

Concerning this corona, we have learned about this from afar. We have never seen a person sick with corona. But we have received teaching on corona. —FGD5, F8

We have learned that it kills mostly politicians and white people. We hope that this disease stays over there, away from us. —FGD3, M6

As another severe viral epidemic, COVID-19 invited comparisons to the Ebola virus disease epidemic that had ravaged the region:

We are very afraid because we have seen the families that lost their family members who died of Ebola. —FGD4, Y1

Ebola killed people, yes, but the radio talks of frightening numbers of deaths due to corona. Really very many. —FGD6, F9

Even with Ebola here, we went to church, to the market, but with corona, no. The churches are closed, and that’s where we go for consolation, imploring God to protect us. But corona closed the churches. It’s serious. —FGD6, F10

Challenges Associated With Hand Hygiene in the Camp

Most FGD participants were aware of the recommendation for frequent handwashing as a prevention measure against COVID-19. However, soap and water were not readily accessible in the camp:

You have to wash your hands. That’s what they say, but we don’t have water here. —FGD4, Y2

Our only source of water is the rain. We collect water when it rains and we keep it. We drink this water. When there is none left, our sisters go to the well to get water. —FGD4, Y2

There is a little stream about 100 m away. That’s what we use for all our needs. —FGD6, F9

To wash our hands, we have water buckets but no soap and it’s not enough because there are only 5 buckets for the whole camp [of approximately 800 people]. —FGD6, F11

They talk about masks, but if we don’t even have soap, how can we ask for more? —FGD2, F2

---

**TABLE 1. Composition of Focus Groups From 3 Internally Displaced Persons Camps in North Kivu, Democratic Republic of the Congo**

<table>
<thead>
<tr>
<th>Focus Group</th>
<th>IDP Participants</th>
<th>Location</th>
<th>Participant Unique Identifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4 men</td>
<td>Mwangaza</td>
<td>M1, M2, M3, M4</td>
</tr>
<tr>
<td>2</td>
<td>4 women</td>
<td>Masosi</td>
<td>F1, F2, F3, F4</td>
</tr>
<tr>
<td>3</td>
<td>4 men</td>
<td>Mwangaza</td>
<td>M5, M6, M7, M8</td>
</tr>
<tr>
<td>4</td>
<td>3 youth (male)</td>
<td>Masosi</td>
<td>Y1, Y2, Y3</td>
</tr>
<tr>
<td>5</td>
<td>4 men</td>
<td>Mwangaza</td>
<td>F5, F6, F7, F8</td>
</tr>
<tr>
<td>6</td>
<td>4 women</td>
<td>Luvangira</td>
<td>F9, F10, F11, F12</td>
</tr>
</tbody>
</table>

Abbreviation: IDP, internally displaced person.
**Impossibility of Physical Distancing in the IDP Camp**

Housing was not conducive to physical distancing for many IDPs. Although many IDPs had individual family dwellings, some were housed in local school buildings, sleeping in classrooms.

*The director and state authorities allowed us to stay here. More and more people came to stay because there was space.* —FGD5, F5

*We don’t pay anything for rent. It’s free.* —FGD1, M1

*In the morning, we move our belongings outside until the end of classes. And at night, we bring back our things into the classrooms we occupy. But since the beginning of corona, we’ve stopped moving things in and out. We keep everything in the rooms where we sleep.* —FGD1, M3

Despite being accommodated by the school, tragically, IDP children did not attend classes:

*We stay with them outside, or else, they come with us to the fields nearby.* —FGD3, M6

*Where are we going to find money to pay the school fees? It’s impossible. We are “wakimbizi” [refugees; those who fled], as they call us.* —FGD3, M5

A repeated theme was the inability to practice physical distancing because of crowded conditions, particularly sleeping quarters in which multiple families occupied a single classroom:

*Here, it’s not possible “ku achana metre moyo moyo” [to stay 1 meter apart; to practice physical distancing]. If it comes here, we will all die. You have seen the conditions we live in. Our room measures 6 m by 5 m, and there are 5 families inside.* —FGD1, M4

*There is no soap, water is a problem, we sleep side by side. Everything is stacked against if this corona arrives here, even if we have, until now, escaped from the “ba chinja chinja” [throat-slitters].* —FGD1, M1

*We are crowded in classrooms like sardines. Isn’t that awful?* —FGD6, F10

One nongovernmental organization came here to educate us about corona. We asked the teacher to give us a practical demonstration. He just smiled! It’s good to teach us, but going back, you should tell the people who sent you that it’s not possible to avoid corona over there. —FGD1, M1

*Do you see how we sleep? During the day, maybe, we can avoid touching each other, but at night we’re in a small room. We’re squeezed one against the other. It’s not possible here.* —FGD2, F4

**Restoring Peace and Security Takes Priority Over Vaccine**

Several respondents were willing to accept vaccination to prevent COVID-19 if a vaccine becomes available:

*I would receive it. For Ebola, people accepted the vaccine.* —FGD5, F5

Others bristled at the idea of a vaccine when more basic needs remain unmet:

*Our concern is safety. Even that vaccine doesn’t matter to us. Let them keep it over there. Even if they vaccinate us, and we continue to live in these conditions, what’s the point?* —FGD5, F5

*If security returns, we will protect ourselves against corona, we will respect all the measures, and it’s only at that time that you can start talking about a vaccine or physical distancing. But in these conditions, I wouldn’t accept this vaccine.* —FGD5, F5

In several FGDs, participants emphasized that COVID-19 prevention recommendations could best be implemented in a more stable, less crowded environment, such as their own homes. Reestablishing security in the region would allow IDPs to return where prevention could be practiced. Other prevention strategies were seen as context inappropriate or even futile:

*The government should bring back peace, we will go back to our homes and we will put into practice all that you have taught us. But it’s impossible to prevent corona here.* —FGD1, M3

*These are measures that don’t apply to us. The only medicine for us here or the only solution that can help us to fight corona here, is security. Bring back peace, and we’ll go back home, where we live in good conditions, and we can respect these recommendations of 1m.* —FGD5, F7

*Me, I’ll only be able to protect myself and my children when I’m at home. We have our own houses with plenty of space, like 6 rooms, but here it’s 1 room. One room with several families. Each has his own activities during the day and you don’t know who will bring you the disease.* —FGD6, F11

**Survey Questionnaire**

Surveys were conducted between 25 and 29 May 2020. One IDP approached declined to participate.
in the questionnaire interview (165/166 [99%] participation rate). Two participants (1 IDP, 1 comparison) had never heard of coronavirus (307/309 [99%] awareness) and were excluded from the subsequent analysis. The final sample consisted of 164 IDPs (66 from Mwangaza; 44 Masosi, and 54 Luvangira) and 143 in the comparison group. There were 74 women (45%) among the IDPs surveyed and 57 women (40%) in the comparison group.

Thirty-five (21%), 82 (50%), and 47 (29%) of IDPs had lived in the camps for less than 1, 1–2, and more than 2 years, respectively. Sixty-six (41%) of families were temporarily sheltered in school buildings. Others lived in structures made from wood, thatch, and mud or brick walls with an iron sheet roof (Table 2). Demographic features, household (family) size, and asset ownership differed significantly between IDPs and the comparison group (Table 2). IDPs surveyed were older, had lower educational attainment, were more commonly farmers, were more commonly married, had a higher median household size, had lower household ownership of indicator assets (radio, cell phone, and bicycle), and had different housing structures than the comparison group (Table 2).

With respect to knowledge of COVID-19, fewer IDPs correctly identified signs and symptoms, and fewer recognized the potential for asymptomatic transmission (Table 3). Overall, 15% of IDPs had sufficient knowledge, versus 30% of the comparison group (OR=0.30; 95% confidence interval [CI]=0.17, 0.53; P<.0001). Other factors associated with low COVID-19 knowledge in bivariate analyses (P<.05) included younger age, larger household size, and lack of radio ownership. In a multivariable logistic regression model adjusting for these possible confounders, IDP status remained statistically significantly associated with lower knowledge (adjusted OR=0.17; 95% CI=0.082, 0.34; P<.0001).

Attitudes and practices toward COVID-19 prevention are shown in Tables 4 and 5, respectively. Despite widespread agreement (89%) that physical distancing was important to prevent COVID-19, a higher proportion of IDPs than individuals in the comparison group reported close contact with someone outside the family in the past 24 hours and a higher proportion had shaken hands with at least 1 person (Table 5).

IDP respondents indicated that movements in and out of the camp were frequent. By self-report, 83 (61%), 62 (38%), and 19 (12%) left the camp on a daily, weekly, and monthly basis, respectively.

In addition, 107 (65%) of IDPs had received a visitor from outside the camp in the past month. Since the pandemic began, IDPs reported leaving the camp less frequently than before in 84 (52%), more than before in 33 (20%) and about the same as before in 46 (28%) of cases.

**DISCUSSION**

Our study is unique among COVID-19 KAP surveys to date for its focus on a displaced population with extreme resource limitations. Other KAP surveys included health care workers or residents of high-income countries with markedly different demographics than our study (e.g., 62% of U.S. and 64% of Chinese participants had a bachelor’s degree or higher, compared to 47% of IDPs in our study who had no formal education at all). Given the radically different challenges of COVID-19 prevention in IDP camps, this study fills a gap in available data from a neglected and isolated population. IDPs differed from neighboring Congolese residents in terms of larger household size (including 46% of families with a member over the age of 60), more extreme poverty, lower educational attainment, less access to information through media and internet, less COVID-19 specific knowledge, lower rate of physical distancing, and reduced access to hand hygiene. These factors, as well as the high mobility of IDPs, leaving and reentering the camp daily for subsistence labor, establish their vulnerability to COVID-19.

**COVID-19 Knowledge**

IDPs and the comparison group both identified local radio as their major source of information on COVID-19 (Table 2). Radio, television, and social media were more common sources of information among the comparison group, whereas church was a more common source among IDPs (Table 2). Other studies in LMICs (Pakistan, Uganda, and Vietnam) showed that health care workers accessed World Health Organization or ministry of health websites (83%–88%), social media (74%–91%), radio or television (46%–79%) for their COVID-19 information, preferences which reflect major differences in education level, employment activities, and access to internet from the IDPs in our study.

Knowledge of COVID-19 was poor in IDPs versus the comparison group (Table 3). Using a similar questionnaire item, 98% and 93% of health care workers in Uganda identified fever and cough as symptoms of COVID-19, compared to 26% and 42% of IDPs, respectively, in our study.
Gastrointestinal symptoms were less frequently identified by both Ugandan health care workers (35%) and IDPs (11%). Misconceptions around COVID-19 transmission (incorrectly endorsing mosquito-borne transmission) were common in both IDPs (54%) and the comparison group (64%) in our study.

Fear of COVID-19 was expressed by 98% of survey respondents, similar to previous observations of high anxiety scores in another survey.
### TABLE 3. Survey Questionnaire Respondents’ Knowledge on COVID-19 Among Internally Displaced Persons, North Kivu, Democratic Republic of the Congo

<table>
<thead>
<tr>
<th>Source of information on COVID-19</th>
<th>Overall (N=307) No. (%)</th>
<th>IDPs (N=164) No. (%)</th>
<th>Comparison (N=143) No. (%)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local radio</td>
<td>278 (90.6)</td>
<td>140 (85.4)</td>
<td>138 (96.5)</td>
<td>.002</td>
</tr>
<tr>
<td>International radio</td>
<td>11 (3.6)</td>
<td>2 (1.2)</td>
<td>9 (6.3)</td>
<td>.03</td>
</tr>
<tr>
<td>Television</td>
<td>12 (3.9)</td>
<td>1 (0.6)</td>
<td>11 (7.7)</td>
<td>.002</td>
</tr>
<tr>
<td>Social media</td>
<td>28 (9.1)</td>
<td>2 (1.2)</td>
<td>26 (18.2)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Church</td>
<td>40 (13.0)</td>
<td>28 (17.1)</td>
<td>12 (8.4)</td>
<td>.04</td>
</tr>
<tr>
<td>Friends</td>
<td>81 (26.4)</td>
<td>50 (30.5)</td>
<td>31 (21.7)</td>
<td>.11</td>
</tr>
<tr>
<td>No response</td>
<td>4 (1.3)</td>
<td>4 (2.4)</td>
<td>0 (0)</td>
<td>.13</td>
</tr>
</tbody>
</table>

**Recognition of illness**

What are the signs and symptoms of COVID-19?^a^  
- **I don’t know**: 62 (20.2) 52 (31.7) 10 (7.0) <.0001  
- **Fever**: 70 (22.8) 43 (26.2) 27 (18.9) .16  
- **Cough**: 171 (55.7) 69 (42.1) 102 (71.3) <.0001  
- **Difficulty breathing**: 109 (35.5) 46 (28.0) 63 (44.1) .005  
- **Sneezing**: 78 (25.4) 41 (25.0) 37 (25.9) .96  
- **Nasal congestion**: 140 (45.6) 59 (36.0) 81 (56.7) .0004  
- **Headache**: 67 (21.8) 46 (28.0) 21 (14.7) .007  
- **Fatigue**: 92 (29.9) 38 (23.2) 54 (37.8) .008  
- **Joint pain**: 80 (26.1) 29 (17.7) 51 (35.7) .0006  
- **Muscle pain**: 27 (8.8) 16 (9.8) 11 (7.7) .66  
- **Loss of appetite**: 14 (4.6) 6 (3.7) 8 (5.6) .59  
- **Diarrhea**: 25 (8.1) 18 (11.0) 7 (4.9) .08  
- **Constipation**: 2 (0.7) 2 (1.2) 0 (0) .50  
- **Bleeding**: 25 (8.1) 14 (8.5) 11 (7.7) .95  

**Asymptomatic spread**

COVID-19 can be transmitted by someone with no symptoms.^b^  
146 (47.6) 60 (36.6) 86 (60.1) <.0001  

**Misconceptions**

COVID-19 can be transmitted by mosquitoes.^b,c^  
125 (40.7) 67 (40.9) 58 (40.6) .19  

COVID-19 can be prevented by eating spicy food.^c^  
21 (6.8) 11 (6.7) 10 (7.0) .90  

**Sufficient knowledge of COVID-19**

Knew key symptoms, did not endorse misconceptions  
79 (25.7) 24 (14.6) 55 (38.5) <.0001  

Abbreviations: COVID-19, coronavirus disease; IDP, internally displaced person.  
^a^ If participant answered “I don’t know,” no further symptoms were solicited. Otherwise, multiple answers were allowed.  
^b^ Used to assess sufficient knowledge of COVID-19.  
^c^ Number (percentage) of participants who erroneously endorsed these incorrect signs, symptoms, or statements.
Many FGD participants considered that COVID-19 was even more severe than Ebola virus disease. From Iran. Surprisingly, many FGD participants considered that COVID-19 was even more severe than Ebola virus disease (in fact, the case fatality rate of Ebola virus disease is more than 60%, compared to less than 2% for COVID-19). Public health messages about the severity of COVID-19 appear to be widely accepted and believed, with FGD participants citing the high number of deaths in wealthy “white” countries and the closing of churches as evidence of danger. Although mistrust in the government (39%), belief in corruption (42%), belief in conspiracy theories (44% and 22%) were prevalent, endorsement of these views did not appear to be associated with prevention practices. This contrasts with surveys of attitudes toward Ebola virus disease in the same area, in which mistrust, rumors, and misinformation were associated with passive and active resistance to control measures.

**COVID-19 Prevention Efforts**

COVID-19 prevention practices vary widely between geographic areas and demographic groups. For example, 98% of Chinese residents at the beginning of the pandemic wore masks when going out compared to 24% of U.S. residents. Mask use was reported by 3.5% of IDPs and 6% of the comparison group, highlighting the lack of personal protective equipment in this setting. Other measures more readily available to IDPs were handwashing (practiced by 98%), distancing from others (48%), and avoiding touching the face (28%), which were reported in proportions similar to the comparison group.

Movement of populations contributes to the spread of COVID-19. In a large refugee camp in Bangladesh, aid workers who enter and leave the camp daily are expected to be the most likely sources of introduction of COVID-19 into the camp. In the IDP camps in our study, the conspicuous lack of aid workers reflects the isolated and hazardous environment, as well as the neglected status of the IDPs. However, 61% of IDPs left the camp on a daily basis, and 65% had received a visitor in the past month. Staying home was practiced less often among IDPs than among the comparison group (P=.039, Table 5).
## TABLE 5. Survey Questionnaire Respondents’ Practices With Respect to COVID-19 Prevention Among Internally Displaced Persons, North Kivu, Democratic Republic of the Congo

<table>
<thead>
<tr>
<th>Prevention practices</th>
<th>Overall (N=307)</th>
<th>IDPs (N=164)</th>
<th>Comparison (N=143)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No. (%)</strong></td>
<td>No. (%)</td>
<td>No. (%)</td>
<td>No. (%)</td>
<td></td>
</tr>
<tr>
<td><strong>In the past 2 weeks, have you done anything to protect yourself from COVID-19?</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>137 (44.6)</td>
<td>77 (47.0)</td>
<td>60 (42.0)</td>
<td>.39</td>
</tr>
<tr>
<td>Yes</td>
<td>168 (54.7)</td>
<td>85 (51.8)</td>
<td>83 (58.0)</td>
<td></td>
</tr>
<tr>
<td>**If so, what?**a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wash hands</td>
<td>149 (88.6)</td>
<td>78 (91.7)</td>
<td>71 (85.5)</td>
<td>.30</td>
</tr>
<tr>
<td>Stay &gt;2 m from others</td>
<td>75 (44.6)</td>
<td>35 (41.2)</td>
<td>40 (48.2)</td>
<td>.45</td>
</tr>
<tr>
<td>Avoid touching face</td>
<td>38 (22.6)</td>
<td>24 (28.2)</td>
<td>14 (16.8)</td>
<td>.11</td>
</tr>
<tr>
<td>Stay home</td>
<td>31 (18.5)</td>
<td>10 (11.8)</td>
<td>21 (25.3)</td>
<td>.04</td>
</tr>
<tr>
<td>Use disinfectant</td>
<td>10 (6.0)</td>
<td>6 (7.1)</td>
<td>4 (4.8)</td>
<td>.75</td>
</tr>
<tr>
<td>Wear mask</td>
<td>8 (4.8)</td>
<td>3 (3.5)</td>
<td>5 (6.0)</td>
<td>.49</td>
</tr>
<tr>
<td>Take medicines without prescription</td>
<td>2 (1.2)</td>
<td>2 (2.4)</td>
<td>0 (0)</td>
<td>.50</td>
</tr>
<tr>
<td>Change diet</td>
<td>1 (0.6)</td>
<td>1 (1.2)</td>
<td>0 (0)</td>
<td>&gt;.99</td>
</tr>
<tr>
<td><strong>Physical distancing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apart from family, have you come in close (&lt;2 m) contact with anyone in the past 24 hours?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>195 (63.5)</td>
<td>115 (70.1)</td>
<td>80 (55.9)</td>
<td>.01</td>
</tr>
<tr>
<td>How many people did you shake hands with in the past 24 hours (not counting family members)?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>155 (50.5)</td>
<td>77 (47.0)</td>
<td>78 (54.5)</td>
<td></td>
</tr>
<tr>
<td>1 to 5</td>
<td>71 (23.1)</td>
<td>31 (18.9)</td>
<td>40 (28.0)</td>
<td>.02</td>
</tr>
<tr>
<td>&gt;5</td>
<td>81 (26.4)</td>
<td>56 (34.1)</td>
<td>25 (17.5)</td>
<td></td>
</tr>
<tr>
<td><strong>Barriers to prevention</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What has prevented you from fully protecting yourself from COVID-19?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of soap</td>
<td>243 (79.2)</td>
<td>150 (91.5)</td>
<td>93 (65.0)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Lack of water</td>
<td>193 (62.9)</td>
<td>110 (67.1)</td>
<td>83 (58.0)</td>
<td>.11</td>
</tr>
<tr>
<td>Insufficient income</td>
<td>67 (21.8)</td>
<td>32 (19.5)</td>
<td>35 (24.5)</td>
<td>.38</td>
</tr>
<tr>
<td>Lack of masks</td>
<td>55 (17.9)</td>
<td>25 (15.2)</td>
<td>30 (21.0)</td>
<td>.26</td>
</tr>
<tr>
<td>Lack of information</td>
<td>51 (16.6)</td>
<td>24 (14.6)</td>
<td>27 (18.9)</td>
<td>.41</td>
</tr>
<tr>
<td>Lack of disinfectant</td>
<td>46 (15.0)</td>
<td>21 (12.8)</td>
<td>25 (17.5)</td>
<td>.34</td>
</tr>
<tr>
<td>Lack of availability of these items</td>
<td>31 (10.1)</td>
<td>18 (11.0)</td>
<td>13 (9.1)</td>
<td>.71</td>
</tr>
<tr>
<td>High prices of these items in the market</td>
<td>46 (15.0)</td>
<td>15 (9.2)</td>
<td>31 (21.7)</td>
<td>.004</td>
</tr>
<tr>
<td>Lack gloves</td>
<td>18 (5.9)</td>
<td>11 (6.7)</td>
<td>7 (4.9)</td>
<td>.66</td>
</tr>
<tr>
<td>I can fully protect myself against COVID-19</td>
<td>28 (9.2)</td>
<td>7 (4.3)</td>
<td>21 (14.7)</td>
<td>.003</td>
</tr>
</tbody>
</table>

Abbreviations: COVID-19, coronavirus disease; IDP, internally displaced person.

a Among respondents who had done something to protect against COVID-19.
Given challenges with hand hygiene and physical distancing in the camps, we speculated that IDPs may have felt disempowered to take small efforts to reduce physical contact with others.

These frequent movements represent opportunities to introduce COVID-19 into the camp. FGD participants explained that daily labor in neighboring fields or trips to the market were imperative to provide for family needs. Thus, unless food security can be assured by other means, restriction of movements to prevent COVID-19 is not viable in the IDP camps studied.

Among IDPs who had taken action to prevent COVID-19, hand hygiene was practiced by 92%. However, the most commonly listed barrier to prevention was lack of soap (92% of IDPs, versus 65% of the comparison group), followed by lack of water (67% of IDPs). Distribution of soap to households in a refugee camp increased handwashing by more than 30% and reduced diarrheal illness in a previous study. In Nigeria, COVID-19 control efforts included sensitization campaigns on handwashing were followed by the distribution of soap to IDPs in Borno State. Inspired by these examples, and responding to the near-universal lack of soap identified in our survey, we included soap distribution in our community feedback efforts.

Avoiding physical contact with others is emphasized as a COVID-19 prevention measure. The majority (89%) of IDPs agreed or strongly agreed that this was an important control measure (Table 4), but 70% had come in close contact with someone other than a family member (versus 56% of the comparison group, \( P = 0.014 \), Table 5). The impossibility of physical distancing in the camp, noted by previous authors, was repeatedly emphasized in FGDs. Sleeping quarters were highly congested, with several families often sleeping in a single classroom. In high-income countries, where shelter-at-home recommendations are more feasible, adherence to physical distancing recommendations remains variable. In the United States, 30% of people reported attending gatherings with more than 50 people (contrary to public health advice), compared to only 3.6% of Chinese survey respondents. In our study, 19% of IDPs had shaken hands with 1–5 people in the past 24 hours, and 34% with more than 5 people, which was statistically higher than the comparison group (\( P = 0.023 \), Table 5). In contrast, 83% of Ugandan health care workers avoided shaking hands due to COVID-19. Given challenges with hand hygiene and physical distancing in the camps, we speculated that IDPs may have felt disempowered to make even small efforts to reduce physical contact with others.

Acceptance of a hypothetical COVID-19 vaccine was high (92%) in a study of Vietnamese health workers. In our FGDs, some participants were willing to accept vaccination as a control strategy, whereas others pointed to futility and inappropriateness of what appeared to them as a stopgap solution, when the overwhelming problem was displacement from their homes.

Expressions of futility or fatalism as expressed by FGD participants in our study are noteworthy and may reflect learned helplessness or loss of self-efficacy among IDPs under extraordinarily difficult living conditions. The theory of learned helplessness describes pessimistic beliefs about the efficacy of one’s actions and the likelihood of obtaining future rewards. The theory has explanatory power among refugees in other contexts, such as risky sexual behavior among victims of sexual or gender-based violence. Similarly, the concept of self-efficacy refers to the degree of externality in control attribution. Low self-efficacy is associated with a fatalistic orientation, as exemplified by a FGD participant’s response. These theoretical frameworks may explain, at least in part, initially puzzling findings such as rejection of a hypothetical vaccine among some FGD respondents and high levels of hand shaking despite awareness and fear of COVID-19.

Limitations
Our study has several limitations. Our survey tool was not validated against a gold standard instrument for the measurement of COVID-19-related KAP among IDPs. However, we took several steps to optimize the validity of the survey: (1) contextually relevant questionnaire items using past surveys from other LMICs and from North Kivu; (2) tacit understanding of the local language and culture by our study team; and (3) implementation of the questionnaire as a verbal interview by local Congolese health workers to allow explanation of questions. The sampling strategy for IDPs and the comparison group was not a fully random sample due to lack of detailed census information. Instead, for IDPs we used geospatial sampling from 3 displacement camps. For the comparison group, we used maximum variation sampling, based on demographic features (age, sex, occupation, and educational attainment). These non-probability sampling methods are widely used, but findings may not be representative of the entire IDP population. Therefore statistical inferences should be interpreted with caution and should be confirmed in studies with a fully random sample of the population of interest (IDPs in North Kivu, DRC). For our primary analysis...
These specific recommendations follow from our observation that leaves no one behind, just as COVID-19 will not respect borders and will not leave behind refugees and IDPs. This result calls for an understanding and appreciation of the complex lives of IDPs and the shared responsibility to ensure COVID-19 prevention measures are in place.

CONCLUSION

In summary, our findings provide a snapshot of the IDP camps as they brace for COVID-19. Awareness and fear of COVID-19 were high among IDPs, but only 15% had comprehensive knowledge of the disease. Significant barriers to implementing COVID-19 prevention measures exist in IDP camps, including crowded sleeping quarters, frequent close contact with non-family members, movement in and out of the camp for work, and lack of access to hand hygiene. Poignantly, IDPs spoke of a desire for peace and a return to their homes, where they could capably prevent COVID-19 themselves. These data from a hard-to-reach population in a zone of insecurity provide a rare glimpse of the desperate conditions under which IDPs survive, leaving them vulnerable to COVID-19. These results call for an ethical, inclusive approach to the global pandemic that leaves no one behind, just as COVID-19 will not respect borders and will not leave behind refugees and IDPs.

RECOMMENDATIONS

These specific recommendations follow from our findings:

- IDPs should be provided with adequate facilities and consumables to implement recommended COVID-19 precautions. These include ample water and soap for hand hygiene and face masks.
- Additional space and housing should be made available to allow IDPs to practice physical distancing, particularly within sleeping quarters. Separate dwellings (e.g., tarpaulin tents) for individual families should be provided. Multiple families sleeping in a classroom (as currently observed) is discouraged.
- Although challenging, restoration of peace by controlling armed conflict in the area is a chief priority for IDPs and would allow a safe return to their ancestral homes where they could more adequately practice COVID-19 prevention.

Acknowledgments: We thank the IDPs who participated in the survey and FGDs. We thank the local authorities, including Kambale Kilwana (bourgmestre adjoint) for providing operational approval to access the IDP camp and conduct the study.

Funding: The Association for Health Innovation in Africa (AFHIA) provided funding for this study.

Competing interests: None declared.

REFERENCES


Global Health: Science and Practice 2020 | Volume 8 | Number 4


En français

La prévention de COVID-19 dans un camp des déplacés internes dans une zone d’insécurité au Nord Kivu, République Démocratique du Congo: une étude avec méthodes mixtes

Message clé


Résumé


Méthodes: Étude avec méthodes mixtes pour la collecte et analyse de données qualitatives (discussions en groupe, DG) et quantitatives (sondage avec questionnaire de 52 éléments).

Résultats: Des DG (23 participants au total) et un sondage (164 DI de trois camps de déplacés et 143 témoins d’un village voisin) ont été organisés en mai, 2020. Les DI étaient statistiquement plus susceptibles d’avoir une plus grande taille de ménage, une pauvreté extrême, un niveau d’éducation inférieur et un accès plus faible à l’information via les médias et l’internet (P <0,05 pour toutes les comparaisons). Les PDI avaient un niveau élevé de sensibilisation (99%) et de peur (98%) du COVID-19, mais des connaissances spécifiques plus faibles (15% de connaissances suffisantes contre 30% parmi les témoins, P=0,0001), une différence qui est restée significative dans un modèle multivariable ajusté pour les effets confondants. Les DI avaient plusieurs défis quant à la mise en œuvre des recommandations pour prévenir le COVID-19. La distanciation physique était impossible dans leurs abris coïncés et 70% des DI ont répondu qu’ils ont été en contact étroit avec une personne autre qu’un membre de la famille au cours des dernières 24 heures (contre 56% des témoins, P=0,014). Les DI devaient souvent sortir du camp pour subvenir à leurs besoins alimentaires, ce qui pourrait permettre l’introduction de COVID-19 dans le camp. 61% des DI sortaient du camp quotidiennement, et 65% avaient eu un visiteur dans le mois précédent. Malgré l’acceptation de l’hygiène des mains pour la prévention, 92% manquaient de savon (contre 65% des témoins, P=0,0001). Les DI cherchaient la paix et un retour au village natal encore plus que d’autres mesures de prévention telles qu’un vaccin contre COVID-19.

Conclusions: Ces résultats fournissent des preuves empiriques soutenant la vulnérabilité des DI au COVID-19 et appellent à l’action pour protéger les populations déplacées négligées.

Peer Reviewed

Received: June 21, 2020; Accepted: September 18, 2020; First published online: December 4, 2020
Meeting the Global Target in Reproductive, Maternal, Newborn, and Child Health Care Services in Low- and Middle-Income Countries


Key Findings

- Progress in reproductive, maternal, newborn, and child health care service coverage is increasing but is uneven between countries and across subgroups (in terms of wealth, place of residence, education, age, and sex) within countries. These coverage gaps are projected to continue.

- By 2030, none of the low- and middle-income countries would be able to achieve the target of universal coverage for oral rehydration therapy for diarrhea treatment or to seek care for acute respiratory infections. Only a few countries are likely to achieve universal coverage for demand for family planning satisfied with modern contraceptive methods, recommended visits for antenatal care, and skilled birth attendant for assistance during birth.

Key Implications

- When designing appropriate interventions for increasing the coverage of reproductive, maternal, newborn, and child health care services, program managers should consider disadvantaged and marginalized populations.

- Acceleration is needed in coordinated global efforts and government policies focusing on marginalized groups, administering cost-effective interventions, and implementing proactive follow-up for routinely scheduled health care services.

ABSTRACT

Introduction: Improving reproductive, maternal, newborn, and child health (RMNCH) care services is imperative for reducing maternal and child mortality. Many low- and middle-income countries (LMICs) are striving to achieve RMNCH-related Sustainable Development Goals (SDGs). We monitored progress, made projections, and calculated the average annual rate of change needed to achieve universal (100%) access of RMNCH service indicators by 2030.

Methods: We extracted Demographic and Health Survey (DHS) data of 75 LMICs to estimate the coverage of RMNCH indicators and composite coverage index (CCI) to measure health system strengths. Bayesian linear regression models were fitted to predict the coverage of indicators and the probability of achieving targets.

Results: The projection analysis included 64 countries with available information for at least 2 DHS rounds. No countries are projected to reach universal CCI by 2030; only Brazil, Cambodia, Colombia, Honduras, Morocco, and Sierra Leone will have more than 90% CCI. None of the LMICs will achieve universal coverage of all RMNCH indicators by 2030, although some may achieve universal coverage for specific services. To meet targets for universal service access by 2030, most LMICs must attain a 2-fold increase in the coverage of indicators from 2019 to 2030. Coverage of RMNCH indicators, the probability of target attainment, and the required rate of increase vary significantly across the spectrum of sociodemographic disadvantages. Most countries with poor historical and current trends for RMNCH coverage are likely to experience a similar scenario in 2030. Countries with lower coverage had higher disparities across the subgroups of wealth, place of residence, and women’s/mother’s education and age; these disparities are projected to persist in 2030.

Conclusion: None of the LMICs will meet the SDG RMNCH 2030 targets without scaling up essential RMNCH interventions, reducing gaps in coverage, and reaching marginalized and disadvantaged populations.
INTRODUCTION

Reducing maternal and child morbidity and mortality and improving reproductive, maternal, newborn, and child health (RMNCH) are top priorities of the global health agenda, particularly for low- and middle-income countries (LMICs). During the era of the Millennium Development Goals (MDGs) between 1990 and 2015, coverage of effective RMNCH interventions to reduce maternal and child morbidity and mortality was scaled up in LMICs. This global initiative showed rapid progress in increasing the coverage of RMNCH care services such as accelerated coverage of demand for family planning satisfied with modern contraceptive methods (mDFPS), presence of a skilled birth attendant (SBA), and radically increased coverage of child vaccinations, while other services had modest progress and a few were far behind in meeting the global targets. Despite significant improvements in health MDGs globally, the population-level inequality between the poorest and richest households and between urban and rural areas did not change in many LMICs. Most importantly, individual-level disparities in terms of gender, age, education, and geographical location suggested further review of global agendas for designing and implementing RMNCH interventions was needed.

In 2015, the United Nations General Assembly summit global developmental agenda shifted from MDGs to Sustainable Development Goals (SDGs). The top priority of SDG target 3.8 is to achieve universal health coverage (UHC), which means that:

-all individuals and communities receive the health services they need without suffering financial hardship.

Forty years after the adoption of the historic Declaration of Alma-Ata, the World Health Organization (WHO) in partnership with the United Nations Children’s Fund (UNICEF) and the Ministry of Health of Kazakhstan hosted the Global Conference on Primary Health Care in October 2018 to recommit to primary health care as the cornerstone of UHC in the new Declaration of Astana. The aims of the declaration are to renew political commitment to primary health care from governments, nongovernmental organizations, professional organizations, academia, and global health and development organizations. RMNCH care services constitute a significant portion of UHC, and reaching and maintaining high rates of coverage of priority interventions indicate the strength of health systems of a country. The results of the Countdown Network suggest that in many LMICs with the highest burden of maternal and child mortality, coverage of some RMNCH care services remains poor, including mDFPS, oral rehydration therapy (ORT), and care seeking for acute respiratory infections (ARI care). However, no projections were made to identify which countries are unlikely to achieve global RMNCH targets. To bridge this evidence gap, the Global Burden of Diseases (GBD) collaborators recently examined trends and projected target attainments of 41 health-related SDG indicators in many countries and territories. Again, projections of these indicators across socioeconomically disadvantaged subgroups are still missing in the existing literature.

Trend analysis helps policy makers and program managers assess current progress, reformulate policies, and design necessary interventions. Projections for RMNCH care services across different socio-demographic dimensions are central to identifying the key priority areas or groups (i.e., identifying the most disadvantaged groups to be covered under interventions) to reinforce or reformulate current policies for achieving country goals. A number of studies, including those conducted by the Countdown Network and GBD, have evaluated the current status, examined trends, and made projections of RMNCH care services and some composite indices at the global, regional, or country level. However, none of these studies captured key interventions for RMNCH separately to make projections across subgroups by sociodemographic stratifications.

In this study, we used the most recent data to assess progress, make projections, and calculate the probability of target attainment and the required average annual rate of change (AARC) for achieving targets of RMNCH care services across various population subgroups within LMICs. We also calculated gaps in coverage of services across a set of sociodemographic dimensions. We did our analyses within and between countries to identify the most disadvantaged countries and groups within countries with inadequate access to RMNCH care services.

METHODS

Data Sources

To calculate the coverage of RMNCH care services, we used macro-level (aggregated) data from large-scale, population-based, nationally representative cross-sectional surveys conducted repeatedly between 1990 and 2018 under the Demographic and Health Surveys (DHS) program in LMICs. Established in 1984 by the United States Agency
for International Development, the DHS program aims to provide decision makers in participating countries with improved information and analyses useful for informed policy choices, improve coordination and partnerships in data collection at the international and country levels, develop the skills and resources necessary to conduct high-quality demographic and health surveys, improve data collection and analysis tools and methodology, and improve the dissemination and utilization of data. The DHS program provides population-based, repeated cross-sectional data that capture a wide range of monitoring and impact evaluation indicators in the areas of population, health, and nutrition. Since the program began, more than 300 nationally representative household-based surveys have been completed under the DHS project in more than 90 countries. Many of the countries have conducted multiple DHS surveys to establish trend data that enable them to gauge progress in their programs. The samples of DHS surveys are generally representative at the national, residence (urban to rural), and regional level (departments, states, or divisions). The collection of the DHS sample is usually based on a stratified multistage cluster design. The data are made available by MEASURE DHS.

DHS obtained data through standardized interviews of women of reproductive age (15–49 years) from the countries under their program, which included a list of prioritized countries for the Countdown cycle. We downloaded, managed, and combined the data from the website to track the progress and make projections about coverage of RMNCH care services at national and subpopulation levels.

**RMNCH Care Service Indicators**

We selected 8 indicators related to RMNCH care services from a range of intervention areas to assess health care systems or delivery for mothers and their children throughout their life stages, across the continuum of care and aligning with global targets. These indicators included mDFPS; antenatal care visits (ANC); presence of an SBA; child immunizations for measles, BCG, and 3 doses of diphtheria-pertussis-tetanus (DPT); ORT for diarrhea treatment; and ARI care. Global standard definitions were used in defining RMNCH care service indicators (Table). Notably, we considered ANC as receiving service at least 4 times from any provider or at least once from a medically trained provider to ensure that the estimates of ANC can be captured from the maximum number of study countries. In addition, we constructed a composite coverage index (CCI) by using the 8 RMNCH care service indicators according to the formula proposed by Boerma et al. The CCI is a weighted mean of the 8 RMNCH care service indicators (Supplement 1 includes more details). To construct the index, we considered all DHS surveys that contained information on all RMNCH care services. However, we performed trend analysis only for the countries with data available for at least 2 DHS rounds to ascertain the trends. The estimates of CCI were not computed for DHS surveys with missing information on any of the RMNCH care services.

**Statistical Analyses**

We estimated the weighted coverage of RMNCH care services as proportions along with 95% confidence intervals from the original survey data. We calculated the coverage of RMNCH care services across subgroups in terms of wealth quintiles, place of residence, education of women/mother, age of women/mother, and sex of child (for child health care services). We used the variables that DHS constructed to present the estimates in the reports. The socioeconomic status of households was determined according to the asset-based wealth index as a proxy measure of household socioeconomic status. The DHS constructed the household wealth index based on household characteristics and ownership of assets by principal component analysis. The households were ranked based on wealth scores and divided into quintiles, from the poorest quintile (lowest 20% of the index) to the richest quintile (highest 20% of the index). The DHS generated variables on place of residence (rural and urban) based on geographical and administrative locations and education (no education, primary, secondary and higher) based on year of schooling. For this study, we categorized the education variable and classified as less than secondary-level education (no education and primary level) and secondary-level or higher education to stratify the study population. See DHS reports for more details. Notably, we restricted our analysis at the country level but not at the regional level for 2 reasons. First, some regions had few numbers of countries and had heterogeneity between survey years, and second, we were interested in assessing progress across individual countries so that country-level programs and policies could be implemented.
To examine trends, Bayesian linear regression models that used a Markov Chain Monte Carlo algorithm of multiple imputations for missing data were applied to estimate the coverage of RMNCH care services and trends from 1990 to 2018 (Supplement 2). We extended this trend analysis to project the coverage of RMNCH care services up to 2030 as set for achieving the SGD target. We reported credible intervals drawn from Bayesian regression analysis along with the estimates. We calculated the probability of achieving the coverage of RMNCH care services as 99% or more by 2030 to understand which countries and populations within each country are on track to achieve universal coverage of these services. We also validated our estimates drawn from regression models with those drawn from the original microdata (Supplement 2 and Supplement 3 Table S12).

We used Stata (version 15.1) and R (version 3.5) statistical software to analyze our data.

### RESULTS

#### Sample Characteristics

We extracted data from a total of 283 surveys from 75 LMICs, of which 64 countries were surveyed at least twice and included in the trend analysis. Projections of CCI were made

| TABLE. Reproductive, Maternal, Newborn, and Child Health Care Services Indicators for the Composite Coverage Index |
|---------------------------------------------------------------|---------------------------------------------------------------|---|
| **Indicators** | **Definitions** | **SDG Target** | **Target Used in This Study for Calculating Probability** |
| Prepregnancy | Demand for family planning satisfied with a modern method among married women | The proportion of married women aged 15–49 years who do not want any more children or want to wait 2 or more years before having another child and are using modern contraception | Universal access | ≥99% |
| Pregnancy | Antenatal care visits | The proportion of women aged 15–49 years in the 3 years preceding the survey who received at least 4 visits from any provider or at least 1 visit from a medically trained provider (i.e., a doctor, nurse, or midwife) during their last pregnancy | Universal access | ≥99% |
| Birth | Skilled attendance at birth | The proportion of livebirths assisted by a skilled health provider (i.e., a doctor, nurse, or midwife) in the 3 years preceding the survey | Universal access | ≥99% |
| Infancy and early childhood | BCG immunization | The proportion of children aged 12–23 months who received 1 dose of the BCG vaccine | Universal access | ≥99% |
| | DPT immunization | The proportion of children aged 12–23 months who received 3 doses of the DPT vaccine | Universal access | ≥99% |
| | Measles immunization | The proportion of children aged 12–23 months vaccinated against measles | Universal access | ≥99% |
| Childhood | Oral rehydration therapy | The proportion of children aged 5 years or younger with diarrhea who received oral rehydration therapy (i.e., oral rehydration salts, recommended home solution, or increased fluids) in the previous 2 weeks | Universal access | ≥99% |
| | Care seeking for symptoms of acute respiratory infections | The proportion of children aged 5 years or younger with symptoms of acute respiratory infections for whom medical treatment was sought from an appropriate health provider in the previous 2 weeks | Universal access | ≥99% |

Abbreviations: BCG, bacille Calmette-Guérin; DPT, diphtheria, pertussis, and tetanus; SDG, Sustainable Development Goal.

*a Universal access is 100%.*
for 59 countries that had information for all 8 RMNCH care services for at least 2 DHS rounds. More than 4.2 million women 15–49 years of age were included for reproductive and maternal health care services, and more than 2.5 million children under 5 years of age were included for newborn and child health care services. A detailed description of the survey year and number of participants are presented in Supplement 3 (Table S1). All the fitted models for projection analysis achieved convergence. The potential scale reduction factor values are summarized in the Supplement 3 (Table S2 to Table S11).

**Trends and Projections**

From 1990 to 2018, the CCI increased in all LMICs and is projected to continue increasing (Figure 1). However, the progressions varied between countries. Based on the current trend, 34 of 59 countries (56.7%) are projected to have less than 80% CCI by 2030. The country-specific projections showed that only Brazil (95.6%), Sierra Leone (93.0%), Cambodia (93.0%), Honduras (90.7%), Colombia (90.5%), and Morocco (90.3%) are likely to have more than 90% CCI. A number of countries (17 of 59 countries) are projected to have poor CCI (less than 70%) in 2030, with the lowest CCI in Guinea (46.7%), Chad (47.1%), Nigeria (48.2%), Yemen (54.6%), and Benin (55.6%) (Figure 2).

Among countries included in the trend analysis, more than 90% coverage is projected to be achieved by 14 of 62 countries for mDFPS, 41 of 64 countries for ANC, 29 of 63 countries for presence of an SBA, 22 of 61 countries for measles immunization, 28 of 60 countries for 3 doses of DPT vaccine, 42 of 61 countries for BCG, 3 of 61 countries for ORT, and 3 of 62 countries for ARI care by 2030. In 2030, the lowest levels of coverage are projected to be in Albania (1.5%) for mDFPS, in Burundi (0.1%) for ANC, in Angola (8.7%) for presence of an SBA, in Kazakhstan (2.4%) for BCG immunization, in Gabon (11.2%) for 3 doses of DPT vaccine, in Nicaragua (7.0%) for measles immunization, in Cameroon (15.3%) for ORT, and in Guinea (14.6%) for ARI care (Supplement 3 Figure S9 to Figure S16).

**Inequalities**

The intracountry inequalities show that significant gaps exist in the coverage of RMNCH care services across population subgroups, and they are projected to continue into the future.

---

**FIGURE 1.** Progress and Projections of Composite Coverage Index in Low- and Middle-Income Countries

Abbreviation: CCI, composite coverage index.
services across population subgroups, and these gaps are projected to continue into the future (Figure 3 and Supplement 3 Figure S17 to Figure S56). The gaps for CCI between the richest and poorest households are projected to be larger, yielding greater CCI among the richest compared to the poorest, with the largest gap in Nigeria by 63.4 percentage points and the smallest gap in...
Peru by 0.5 percentage point (Figure 4). In contrast, the CCI is projected to be greater among the poorest compared to the richest by 23.9 percentage points in Liberia. Most of the countries with the largest richest-poorest gaps are likely to experience larger urban-rural gaps as well in the CCI, with the greatest CCI gap in the urban population by 25.1 percentage points in Nigeria and the smallest gap in Guatemala by almost nil (Figure 4). In line with richest-poorest and urban-rural gaps, the coverage gaps between women with less than secondary-level education and women with secondary-level or higher education are also expected to remain larger in 2030, with the largest CCI gap among the women with secondary-level education or higher compared with women with less than secondary-level education in Nigeria by 36.1 percentage points and smallest gap in Indonesia by 0.1 percentage points. The CCI gaps between adolescent and adult women are also apparent, but these gaps are considerably narrower than gaps observed across wealth, residence, and education (Figure 4). Indicator-specific projections highlight that the gaps in the coverage of all 8 RMNCH care services are expected to be largely apparent in 2030, predominantly between the richest and poorest at the national level and across urban-rural residence (Supplement 3 Figure S57 to Figure S88).

We also tracked progress in newborn and child health care services based on sex of the child. By 2030, the projected coverage of ORT will be less than 80% in most of the LMICs for both boys and girls (Supplement 3 Figure S92). Similarly, the coverage of ARI care for both boys and girls is projected to be less than 80% by 2030 in most of the countries (Supplement 3 Figure S93). The current sex-based gaps in child immunization rates are also likely to persist in some countries in 2030 (Supplement 3 Figure S89 to Figure S91).

Probability of Target Attainment
According to the posterior probability, Brazil (72%) has the highest probability of achieving universal CCI, followed by Kazakhstan (40%) and Sierra Leone (20%) (Supplement 3 Table S13). Our results indicate that it is unlikely that any of the LMICs will achieve universal coverage for certain RMNCH care services.

None of the LMICs are likely to achieve universal CCI by 2030, although some may achieve universal coverage for certain RMNCH care services.

Global Health: Science and Practice 2020 | Volume 8 | Number 4

FIGURE 3. Trends in Predicted Composite Coverage Index Across Countries by Wealth Quintiles
in Armenia, Brazil, Cambodia, and Jordan. But the probability of achieving universal coverage for other services is close to zero for the majority of the countries. The posterior probability of achieving universal coverage of RMNCH care services across subgroups is also zero for most of the countries (Supplement 3 Table S15 to Table S22). Additionally, we calculated the posterior probability of countries achieving at least 75% coverage for mDFPS. The results showed that nearly one-third (19 of 62 countries) of the countries are on track to achieve the target of at least 75% mDFPS coverage with at least 90% probability of attaining the goal (Supplement 3 Table S23).

**Change Rates**

The progression rates in CCI varied over time; slower rates of progression in CCI are projected in most of the countries during 2019–2030 compared with the progression rate during 1990–2018 (Figure 5). Some countries (e.g., Maldives, −0.2%) had retrogression in CCI during 1990–2018 that will continue during 2019–2030. The calculated AARC shows that achieving the target will require ramping up the rate at which CCI increases annually between 2019 and 2030, particularly by 9.5% in Chad, 7.5% in Nigeria, 7.2% in Guinea, and 6.8% in Yemen (Figure 5). The largest improvements are required for mDFPS for most of the countries, urgently in Albania by 28.2%, Maldives by 15.0%, Democratic Republic of the Congo by 13.3%, Chad by 13.2%, and Yemen by 11.1% (Supplement 3 Table S28). Acceleration in improving the coverage of both ORT and ARI care needs to be at an annual rate of 3%–10% for almost all the countries to achieve the targets (Supplement 3 Table S61 and Table S67). However, the AARC varied across different sociodemographic dimensions within countries (Supplement 3 Table S24 to Table S72 includes details for all RMNCH care services).

**DISCUSSION**

This study provides the most up-to-date estimates on the progress of LMICs toward the key RMNCH care services, and it predicts coverage of these services by 2030 to detect whether RMNCH targets can be achieved. Based on current trends, we demonstrated that none of the LMICs would be
able to meet the target coverage for either ORT for diarrhea treatment or ARI care. Although the coverage of RMNCH care services is increasing, the coverage gaps across sociodemographic dimensions remain and are projected to persist. Substantial variations exist in the coverage of RMNCH care services between countries and between subgroup levels within countries. These results emphasize the need for effective policies focusing on marginalized groups, administering cost-effective interventions, and implementing proactive follow-up for routinely scheduled health care visits to ensure universal access to RMNCH care services. The results of this study provide evidence to inform global and country leaders and policy makers about the country-specific situations at national and subgroup levels and highlights key areas of interventions (such as improving ORT and ARI care services) that need urgent attention for increasing the coverage of these services through allocating national funding and resources toward achieving the 2030 target for RMNCH care services.

Specific Services
Our results indicate that all countries are unlikely to achieve universal CCI. Some countries are on track to achieve universal coverage for childhood immunization for BCG, DPT, and measles vaccines. Concurrently, some countries such as Maldives, Nigeria, Tajikistan, Yemen, Chad, and Zimbabwe are projected to have less than 80% childhood immunization coverage in 2030. The results of our study demonstrate that coverage of 2 care-seeking services for child morbidity, ORT, and ARI care will be remarkably lower (less than 50% in 25 countries of 61 for ORT and 18 countries of 62 for ARI care) than the target coverage in LMICs. The probability of achieving universal coverage for these 2 services by 2030 is roughly zero for all countries, except Sierra Leone (57% probability) for ORT and Brazil (39% probability) for ARI care.

By 2030, universal coverage is expected to be achieved by Liberia for mDFFPS; Maldives, Armenia, and Cambodia for ANC; and Armenia,
Honduras, and Jordan for presence of an SBA. However, our results demonstrate that most of the countries are struggling to achieve universal coverage of mDFPS, ANC, and presence of an SBA. In addition, the target coverage of these 3 services will not be achieved by most of the subgroups within each of the LMICs. The lower coverage of mDFPS, ANC, and presence of an SBA among the poorest populations, those living in rural areas, and women with less education will impede LMICs, particularly countries in South and Southeast Asia and sub-Saharan Africa, in achieving the target coverage for these 3 services. Although the overall CCI increased, we project that LMICs and all subgroups within LMICs will not be able to reach universal CCI by 2030, especially due to the lower CCI led by mDFPS, ORT, and ARI care among adolescent girls and mothers and among women and mothers who are poor, have less education, and live in rural areas. Our findings correspond with those from previous studies, with negligible variations, which were mainly driven by the number of time points with available data analyzed.

Equity
Based on our results, large coverage gaps exist in childhood immunization coverage between poor and rich households, rural and urban populations, mothers with low and high education levels, and adolescent and adult mothers. To achieve universal immunization coverage by 2030, most countries need to further ramp up of coverage, particularly for the poorest and rural populations and less educated and adolescent mothers in LMICs with low coverage of RMNCH care services. To increase the coverage of RMNCH care services, equitable, appropriate, and focused programs need to be implemented, and resources need to be allocated to increase availability, accessibility, and use of services, particularly for those groups shown to be the furthest behind in the current study (such as poorest, rural, and less educated populations). These programs may help countries to reduce coverage gaps within countries toward achieving the global target of UHC.

Our analysis found considerable disparities in the coverage in ORT and ARI care in terms of wealth, place of residence, education and age of mother, and sex of the child. These gaps may persist until 2030 in some LMICs, predominantly in countries in sub-Saharan Africa. In most LMICs, the coverage for ORT and ARI care will be less than 80% across most subgroups. This projection may partly be explained by broader baseline gaps in ORT and ARI care among the subgroups in LMICs. In general, people who were poorest, residing in rural areas, or were adolescent and less educated mothers will remain vulnerable for achieving the target coverage by 2030. This finding suggests that children belonging to either of these vulnerable groups should be given special consideration in the design of interventions to scale up RMNCH care services.

The gap in CCI must be considered before planning for actions to improve the strengths of health systems. As the projected estimates reveal that none of the LMICs will be able to achieve the CCI target by 2030, we postulate that the lower CCI among the poorest, rural, women/mother with less than secondary-level education, and adolescent women/mother groups has a substantial contribution to the lower CCI. To achieve universal coverage, accelerations on improvements are essential in LMICs with nearly 4% improvements in annual national coverage and 2%-5% improvements in annual coverage at subgroup levels in LMICs. All countries are projected to fail to achieve the CCI target coverage by 2030 at national and subgroup levels, and only some Latin American and Caribbean countries will have more than 80% CCI and are on track for achieving the target if effective RMNCH strategies can be implemented. However, most sub-Saharan African countries will be far behind in reaching the CCI target. Similar to LMICs, the subgroup coverage gaps in RMNCH care services will constitute the key driver behind this target failure. To accomplish the goals of achieving universal access to RMNCH care services, sub-Saharan African countries need to increase the coverage of RMNCH care services by more than 3 times during 2019–2030 than what was calculated during 1990–2018, giving particular attention to the poorest, rural, and less educated and adolescent women/mothers.

For the future progress of RMNCH care services, it is imperative to understand the reasons for lower coverage or gaps in coverage and the associated factors for high or low coverage across different geographical settings. It is well known that between- and within-country inequalities and the lack of financial resources are major constraints for improving RMNCH. In line with previous evidence, our study also demonstrates that coverage of health care services that can be scheduled in advance, such as immunization coverage, were higher and are likely to be achieved by 2030, while those that require emergency on-demand availability of workforce and specialized equipment...
Strengths and Limitations
In this study, we used globally recognized nationally representative data to calculate the coverage of RMNCH care services that provided reliable estimates of trends along with the AARC during different periods. We used a set of globally accepted standard outcome interventions that cover life stages of women during prepregnancy to child­hood of their offspring at the population level and across the continuum of care. The use of large samples from population-based household surveys enabled us to estimate national and subgroup-level trends across countries as well as across subgroups within countries. The unique survey methodology and measurement of the DHS allowed this study to make cross-country comparison of estimates as well. However, the findings of our study need to be interpreted in light of some limitations.

For cross-country comparison, we considered a doctor, nurse, or midwife as skilled personnel for assisting birth as recommended.21 This underestimates the coverage estimates of skilled birth attendance for some countries that may have other skilled service providers, such as paramedics, family welfare visitors, and community skilled birth attendants. Because some countries had too few surveys with available information, we could not make projections of RMNCH care services for those individual countries. Interventions to improve RMNCH care services come in phases and may reach some subpopulations before others. However, we were unable to examine whether the past changes would proceed uniformly in the future within and across countries due to the heterogeneity in survey years within and across countries. Fewer data points for some RMNCH care services for some countries may have created wider credible intervals for the projected estimates of the service coverage (e.g., for CCI in Nicaragua). Credible intervals with a wide range are normal for projection analysis, but they could be narrowed by having multiple time points available (e.g., for CCI in Bangladesh). Calculating more realistic probability estimates is also possible with wider credible intervals. Estimates drawn from representative data collected from multiple sources may better project the future directions of the RMNCH care services with lower uncertainty.

CONCLUSIONS
Although the coverage of RMNCH care services is improving in LMICs, the progress is uneven within and between countries and insufficient to meet the health SDGs. Most sub-Saharan African and South and Southeast Asian countries are very unlikely to achieve target coverages by 2030 due to low coverage overall and high coverage gaps in RMNCH services between the richest and poorest, urban and rural, and high and low education subgroups. These results reflect the urgent need for health interventions targeting disadvantaged countries and their subgroups to achieve universal access to health services and to reduce health inequalities during the SDG era. Increasing funding for RMNCH care through cost-effective interventions may strengthen health care services and can help interventions reach marginalized and disadvantaged people. Country leaders, stakeholders, and agencies need to undertake multidisciplinary collaborative actions by going beyond their commitment in allocating resources, implementing programs, and monitoring the progress and gaps in RMNCH care services toward achieving SDG target 3.8 by 2030.

Acknowledgments: We thank the DHS program for providing access to DHS datasets.

Funding: MMH acknowledges the financial assistance from the University of Queensland and the Commonwealth Government of Australia to undertake his PhD.

Competing interests: None declared.

REFERENCES


Peer Reviewed

Received: March 1, 2020; Accepted: August 26, 2020; First published online: October 8, 2020


© Hasan 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-20-00097
Contraceptive Method Mix: Updates and Implications

Jane T. Bertrand,a John Ross,b Tara M. Sullivan,c Karen Hardee,d James D. Shelton*e

Key Messages

- Contraceptive method mix reflects both supply and demand.
- Recent trends include a progression in hormonal methods toward implants in sub-Saharan Africa, and where HIV is common, more condom use in some countries.
- However, dominance of 1 method in the mix remains very common, though countries and regions throughout the world are diverse as to which method is dominant.
- Our analysis argues for continued concerted efforts of programs to increase contraceptive method choice.
- There is no ideal method mix; client preferences are key.

ABSTRACT

Context: Improving contraceptive method choice is a goal of international family planning. Method mix—the percentage distribution of total contraceptive use across various methods—reflects both supply (availability of affordable methods) and demand (client preferences). We analyze changes in method mix, regional contrasts, and the relationship of the mix to contraceptive prevalence.

Methods: We use 789 national surveys from the 1960s through 2019, from 113 developing countries with at least 1 million people and with data on use of 8 contraceptive methods. Two measures assess the “evenness” of the mix: method skew (more than 50% use is by 1 method), and the average deviation (AD) of the 8 methods’ shares from their mean value. Population weighted and unweighted results are compared because they can differ substantially.

Results: Use of traditional methods has declined but still represents 11% of all use (population weighted) or 17% (unweighted country average). Vasectomy’s share was historically low with the exception of a few countries but is now even lower. The previous trend toward greater overall evenness in the mix has slowed recently. Sub-Saharan Africa shows a hormonal method progression from oral contraceptives to injectables to implants in a substantial number of countries. In some countries with high HIV prevalence, the condom share has increased. The leading method’s share differs by region: female sterilization in Asia (39%) and in Latin America (31%), the pill in the Middle East/North Africa (32%), and the injectable in sub-Saharan Africa (36%). Method skew persists in 30% of countries. “Evenness” of mix is not related to contraceptive prevalence.

Conclusion: The marked diversity in predominant methods underscores the conclusion that no single method mix is ideal or appropriate everywhere. But that diversity across countries, coupled with the persisting high degree of extreme skewness in many of them, argues for continued concerted efforts for programs to increase method choice.

INTRODUCTION

A key principle in both quality of care and the broader rights-based approach to family planning is method choice. As defined by the U.S. Agency for International Development, method choice exists when:

client-centered information, counseling, and services enable women, youth, men, and couples to decide and freely choose a contraceptive method that best meets their reproductive desires and lifestyle, while balancing other considerations important to safety, correct use, or switching methods.
Method choice is a guide for optimal delivery of family planning services. To help ensure that clients’ needs are met across time and changing circumstances, the World Health Organization in 2014 recommended that family planning programs include at least 5 types of modern contraceptive methods: barrier, short-term reversible, long-term reversible, and permanent, along with emergency contraception. Method mix is an indicator that shows the pattern of actual use. It gives the percentage distribution of use across all methods in a given country, also known as “method share.” It can be calculated either in relation to women married/in-union or to all women of reproductive age, using data from a population-based survey.

Interest in the method mix of contraceptive use goes back at least to the 1980s and early 1990s, focusing not just on the empirical patterns but also upon what might constitute an “appropriate” mix. Choe and Bulatao (1992) compared methods for finding an appropriate mix, based partly on the life stage of the woman, whether before or after marriage, between births, or after the final birth. Following that, Galway and Stover (1995) published a tool online to help calculate an appropriate mix, based on users’ personal profiles, the prevailing mix, method preferences shown in surveys, and method characteristics, using Kenya as a case study.

Potter (1999) argued that some mixes could become outdated as not fitting the emerging needs of the population. That could occur when the early pattern of contraceptive supply and use persisted due to being reinforced by feedback from users and program managers, as illustrated in case studies from Brazil and Mexico.

Subsequently, Bertrand et al. directed attention to method mix in which a single method accounted for more than 50% of all use (a “skewed” mix) and its relation to the quality of a national family planning program. Related analyses with data sets covering most developing countries followed, giving attention to changing mix patterns and their relationship to socioeconomic correlates and to the efforts of family planning programs. Ross et al. developed a different approach; rather than looking at the skew due to a single method, it took account of the distortions in mixes across all methods: the average deviation (AD) method, which is employed below along with measures of skew.

An historic disturbance to the prevailing method mixes occurred especially in countries in east and southern sub-Saharan Africa due to the steep rise in use of the injectable method starting in the 1990–1995 period. Several analyses were conducted to trace these changes in the context of their effects upon other methods. Rossier and Corker reviewed the use of traditional methods in sub-Saharan Africa. Rossier and colleagues also documented the underreporting of traditional methods that can occur in surveys. Recently, the United Nations (UN) Population Division published a global review of use by method, for all women rather than married/in-union women, and with regional averages population weighted.

To the extent that method choice (defined above) is an underlying principle of quality family planning service delivery in developing countries, it has important implications for an “ideal” method mix. In contrast to earlier attempts to identify an “appropriate” method mix for a population, one can argue that the “ideal” method mix occurs when all women in a given country are using their desired method, consistent with the conditions outlined for convenient method choice. However, we are unaware of any research that has attempted to measure method mix from this perspective.

Method mix reflects both supply and demand. On the supply side, method choice is optimized when the full range of contraceptives is available with close geographic access, with no stock-outs or cost barriers, with adequate counseling on the methods and on the management of side effects, and with freedom from any provider bias toward or away from particular methods. Method skew may signal that potential users have only a limited choice, based on shortcomings in the supply environment. However, the measure of skew, by itself, provides little insight into the reasons for the constraints on choice.

Method mix is also influenced by demand, including individual or societal preferences. Clients’ attitudes are subject to many influences. They may seek a method because it dominates the environment of what is available in the national program, as with sterilization in India. The introduction of a new method with low cost may stimulate a demand for it, as with the implant in numerous countries in sub-Saharan Africa. Demand for a given method can be adversely affected by known side effects, health concerns, misconceptions, and rumors. Donors may influence the supply of methods by decreasing the cost and supporting training in the provision of the method (e.g., implants). Program directors and providers may also emphasize certain methods over others. The private sector can also influence the availability of methods. Cultural influences are important. They inhibit sterilization use in the Middle...
We present new evidence on patterns and trends in the method mix to provide the most current picture available.

East partly on religious grounds; Islam, as practiced in some countries, equates sterilization with prohibited mutilation of the body. By contrast, the widespread use of female sterilization in Latin America is accompanied by societal acceptance of the method as a practical means of controlling further childbearing among women who achieve their desired family size at a young age. Women may especially dislike methods, such as the intrauterine device (IUD), that require pelvic examinations. Also, for unmarried young women in some societies, confidentiality of contraceptive adoption, combined with private practice without partner or family interference, is important to avoid stigma.

Total demand for contraception (influenced by the desired family size), as well as the method-specific demand, interact with and are mediated by the constraints in the supply environment. Finally, the relative significance of supply and demand factors on method use varies across countries and across subnational entities. All of this reminds us that a perfect method has yet to appear nor can any 1 method ever be expected to be right for all clients.

This article presents new evidence on patterns and trends in method mix, overall and by regions, as well as in selected countries, for married/in-union women of reproductive age. Overall, our aim is to provide the most current picture available but with some historical information and the entire time trend for 2 illustrative countries.

The objectives of the article are to:

1. Document recent changes in contraceptive method mix in developing countries
2. Examine the dominance of specific methods by region and by country
3. Test the relationship between evenness of method mix and contraceptive prevalence
4. Explore the implications of method skew for program applications

■ DATA AND METHODS

Data for this article come from a large compilation of national surveys prepared by the UN Population Division (UN Department of Economic and Social Affairs) in its 2019 release. The database contained 1,202 surveys, from which we retained 789, using the following criteria: the country is (1) classified by the UN definition as being in the developing world, (2) has a population exceeding 1 million, and (3) has the necessary information for contraceptive use of 8 methods: female sterilization, male sterilization (vasectomy), IUD, implant, pill, injectable, condom, and traditional methods; these 8 are the focus of the analysis. Other methods in the UN series, such as the female condom, Lactational Amenorrhea Method (LAM), vaginal barrier, and emergency contraception, appear infrequently or at zero levels in the UN compilation of surveys. Moreover, the focus on these 8 methods provides continuity with earlier publications.

Although family planning programs and donor agencies promote modern methods of contraception, we have kept traditional methods in this analysis because its use persists in numerous countries. Also, it allows us to assess the evolution in method mix from traditional methods to modern methods (or vice versa, if that is occurring).

Half of the surveys are either Demographic and Health Surveys (DHS) (34%) or Multiple Indicator Cluster Surveys (MICS) (16%), and another 27% are listed as “national surveys” done by various agencies. The rest consist of the Contraceptive Prevalence Surveys (CPS) or Reproductive Health Surveys (RHS), largely from Latin America; the Pan Arab Project for Child Development Survey and Pan Arab Project for Family Health Survey, mainly in the Middle East; and the Performance Monitoring and Accountability 2020 (PMA2020) Surveys from 11 countries.

By region, 24 countries are in Asia (including 5 in the Central Asian Republics), 23 in Latin America, 21 in the Middle East/North Africa, and 45 in sub-Saharan Africa, totaling 113. The numbers of surveys in these regions, respectively, are 223 from Asia (with 20 in the Central Asian Republics), 160 (Latin America), 120 (Middle East/North Africa), and 286 (sub-Saharan Africa), totaling 789.

Regarding timing, the 789 surveys occurred from 1963 to 2018; the median survey date was 2001. By decade, the percentages were 1960s (0.6%), 1970s (7%), 1980s (14%), 1990s (22%), 2000s (31%), and the 2010s (25%). For just the latest surveys in the 113 countries, most occurred in recent years, 51 between 2010 and 2014 and 45 between 2015 and 2017. Only 17 were conducted before 2010. For analyses across time, we have annualized the trend within each country, and in analyses of regional trends we have weighted the data by population size. We have not adjusted the regional comparisons for calendar time; the dates between...
the earliest and latest surveys in 1 country are not necessarily the same as in other countries; moreover, the surveys can occur at different periods during the development of the national family planning program. Finally, the earliest-latest survey comparisons give the long-term picture of change, and they avoid comparisons between surveys occurring close to each other, which can introduce atypical short-term fluctuations. Correlational analyses showed that there is essentially no relationship between the size of the gap between the earliest and latest surveys and the pace of annual changes.

In this type of cross-national analysis, one can present the data as weighted (based on the population size of each country) or unweighted (in which each country has equal weight). Both have their place. Weighted data—which give every person equal importance—are useful, for example, in calculating the number of modern contraceptive users in the 69 poorest countries in the world monitored by FP2020. These estimates appropriately reflect the disproportionate contribution of large countries. By contrast, unweighted data—which give every country equal importance—are useful in assessing progress by country, as in the case of the UN Sustainable Development Goals. Rather than choose between weighted or unweighted data, we have opted to present both in this article.

To assess mix, we employ 2 indicators. The first is “method skew,” which indicates whether any single method accounts for more than half of all contraceptive use. When that extreme share occurs, the other 7 methods are necessarily relegated to smaller shares, well below 50%. Other rules could be used (e.g., 60% in the FP2020 reports), but to be consistent with previous articles on method skew, we have retained the cutoff point at 50.

The second measure is the AD, which Ross et al. (2015) introduced to capture the evenness of the mix across all methods, thereby augmenting the information on skew by a single method. Since use of the 8 methods adds to 100%, the average of the 8 shares is always 12.5%, and the share of each method varies around that average. The AD measure looks at the average of the deviations to capture the spread of the shares. A large spread usually indicates that just 1 or 2 methods account for most contraceptive use and the others rather little. That again suggests a limited choice.

In general, the closer each method is to the mean of 12.5%, the lower the AD value. Over time, if 1 method’s share moves closer to the mean, either from above or below it, that reduces the AD value. Depending upon the country, certain methods may take zero values in an early survey if they are severely neglected or not yet made available; that makes for a high AD value. On the other hand, the introduction of a new method can increase its share of the mix, moving it up from a zero share toward the mean of 12.5%. That would result in a decline in the AD value.

If all 8 methods had an equal share of the mix, at 12.5% each, the AD value would be zero; in practice that has never occurred. The actual AD values range from 5 to 19. Perfect evenness does not exist in any country, nor would family planning experts expect it to. Further, no AD value should be considered the “ideal”; it simply serves as an objective measure that allows one to assign a score of evenness or “balance” to the method mix of each country.

In the following sections, most averages are population weighted. The levels and changes in the mix are first calculated for each country and then averaged to obtain regional estimates.

The analysis includes the following specifics:

- For trends, we calculated the change in method mix between the earliest survey and the most recent survey conducted in each country and then determined the average change for each region.
- For the latest levels, we determined the contraceptive method mix for each region and for all countries using the most recent survey conducted in each country.
- We illustrated the long-term dynamics for changes in method mix for the 2 examples of Rwanda (1983–2015) and Ghana (1979–2013).
- We identified the 34 countries with a method skew (>50%) as of the most recent survey along with the method causing the skew.
- We obtained the distribution of countries by the AD value and examined its relationship to the maximum share of use by any method, based on the most recent surveys in all countries.
• We determined the relationship between the AD value and the contraceptive prevalence rate (CPR), based on the most recent surveys in all countries.

RESULTS

This analysis captures the dramatic changes in method mix over several decades of international family planning. Among the 113 countries studied, 109 had 2 or more surveys, allowing for changes between the earliest and the latest surveys (Supplement). The time periods varied around an average interval between surveys of 17 years. Figure 1 summarizes these changes by region and for all countries. The changes are annualized to allow for dissimilar observation periods, and they are population weighted. The bars above the line denote gains by a method; those below the line, losses. Changes within each region add to zero.

For all countries, traditional methods lost an annual average of 0.42 points of share, or 4.2 points over 10 years. The pill also lost share, and small losses occurred for male sterilization and the IUD. Meanwhile, female sterilization, the implant, the injectable, and the condom gained shares.

Among regions, Asia showed the smallest changes while sub-Saharan Africa showed the most, with Latin America and the Middle East/North Africa experiencing intermediate degrees of change. The most extreme shift was in sub-Saharan Africa with the injectable replacing traditional methods. In the early years, its CPR was often low, so that traditional methods could represent a large percentage of a small pie.

As explained in the Methods and Data section, we addressed any concern about methodological differences across the survey types by rerunning the results just with the DHS and MICS surveys and found essentially no differences in the main levels and patterns. We therefore decided to use the full set of surveys to enlarge the base by regions and to augment the time trends.

Key Changes in Method Mix in Recent Years

From this analysis, we identified 4 key trends.

1. Traditional Method Use Has Declined Over Time but Remains Substantial

Traditional method use remains perplexing and somewhat controversial among international family planning experts. Some argue that programs should actively try to move clients from traditional to modern methods, given the greater effectiveness of the latter in preventing pregnancy. Others contend that traditional methods, which are “natural,” serve a valuable purpose; they are noninvasive,
free, always available, and have no side effects. Some maintain that while family planning programs should not necessarily promote traditional methods, people should know how to use them (particularly withdrawal) in case they are having sex without any other method available. Still others view traditional methods as a bridge to modern contraceptive use, especially when a woman has experienced an unplanned pregnancy while using a traditional method.

Despite the tremendous strides made in family planning programming worldwide over the past 5 decades, a surprising 11% of all users, or about 1 in 10, continue to rely on traditional methods. In each country, trends in the use of each method are derived from the change between the earliest available survey to the latest one. This approach provides the experience of the country over the long term, while mitigating short-term fluctuations and measurement errors. The annual rate of change is used to allow for different observation periods between the surveys.

Averaging over all countries, the annual rate of decline for traditional methods has been 0.42%, or 4.2% over 10 years (Figure 1, total bar). Regional averages varied considerably, as the above examples suggest. The loss of traditional share was least in the Middle East/North Africa at only 0.10% and greatest in sub-Saharan Africa at 1.42%. The loss was quite different between Asia (a low 0.16%) and Latin America (a much higher 0.62%). Thus, the loss of share for traditional methods was considerable and quite variable by region. The large loss in sub-Saharan Africa probably reflects the high initial reliance on traditional methods, falling to lower levels as modern methods rose.

Two country examples vividly illustrate the possible changes in method mix over time. In Rwanda, the traditional share fell from 92% in 1983 to only 11% in 2015 (Figure 2), a decline of 81%, the largest on record. In Ghana, (Figure 3) the traditional share fell from 52% to 18%, a

**FIGURE 2. Rwanda: Changes in Method Mix Between 1983 and 2014**

Despite tremendous strides in promoting modern contraceptive methods over the last 50 years, about 1 in 10 users still rely on traditional methods.
34% decline, less than in Rwanda but down to only one-third of the starting level.

2. Vasectomy’s Share of Method Mix Has Declined, From Low to Lower
Vasectomy has had limited uptake for a combination of reasons related to supply and demand, especially in recent years. In the 789 surveys examined here, vasectomy equaled or exceeded the “equal share” of 12.5% only in the Republic of Korea (all surveys 1985–2006), Nepal (all surveys 1976–2011), and Thailand (14.2% in 1969), though close to equality in China (12.1% in 1992). Relatively high values elsewhere occurred mainly in the early days of family planning programming, from the 1960s through the mid-1970s, when few other methods were available. As with traditional methods, in early programs, the percentages for vasectomy often represented a large share of quite low prevalence.

Vasectomy’s share has undergone a drastic decline in 7 countries where it was important, between the peak year of its use and the year of the most recent survey. In each country, its share of the method mix has plummeted. Here are the declines, in order of the starting levels of the shares: Nepal 47.1 to 10.5 (the highest current figure), Thailand 14.2 to 0.5; China 12.1 to 1.7; Myanmar 10.7 to 0.6; India: 8.6 to 0.6; Sri Lanka: 8.2 to 0.0; and Bangladesh 7.1 to 1.9. Most other countries in the data set showed small, non-zero percentages for vasectomy, and in no country did vasectomy increase its share over time.

Regarding national policies, a few countries have promoted the voluntary use of vasectomy with some success (for example reaching 5% of the method mix in Colombia by 2016 and in Brazil by 2013), but the method faces cultural and gender barriers, especially in sub-Saharan Africa, with concerns that men will lose their strength and masculinity if they have the procedure.31 Vasectomy also faces religious barriers in Muslim countries,32,33 as does female sterilization in most Muslim countries. However, female sterilization accounts for a quarter of all use in Pakistan, about 7% of the mix in Bangladesh, 13% in Turkey, and 18% in Iran. In any case, few programs have opted to promote vasectomy in recent years, and in practice, policy makers have shown little political will to explicitly promote vasectomy.

Modifications in the mix reflect the relative changes in the prevalence of the methods over time. If, for example, the use of traditional methods

---

**FIGURE 3. Ghana: Changes in Method Mix Between 1979 and 2017**

[Graph showing changes in method mix from 1979 to 2017 with a legend indicating the different methods: Female Sterilization, Male Sterilization, IUD, Implant, Injectable, Pill, Condom, Traditional Methods.]

Abbreviation: IUD, intrauterine device.
remains about the same while the use of modern methods increases, leading to a rise in total contraceptive use, that produces a diminishing share of all use for traditional methods. In India, total prevalence of use rose from 40.7% to 53.5% between the 1992/93 and the 2015 surveys. Female sterilization rose from 27.4% to 36.0%, while male sterilization declined from 3.5% to 0.3%. For the mix, that translates to a stable female share of 67.3% in both surveys and a decline in the male share from 8.6% to 0.6%.

For prevalence, overall sterilization was gaining. Vasectomy was declining, but female sterilization was increasing enough to more than compensate, and it was doing so in the context of other method changes (Figure 1). For shares that was the general pattern: in a full set of within-country comparisons, the share for female sterilization rose on average twice as fast as the male share did.

**Weighted vs. unweighted results:** The mix looks quite different when the results are weighted by the population size of the country versus unweighted, when each country has an equal weight.

Overall, in Table 1, last row, 11% of all users rely on traditional methods (weighted data), whereas the country average is higher at 17%.

The difference reflects the impact of the largest countries, where fewer rely on traditional methods. Other methods also reflect the impact of the largest countries. In Asia, 39% of all users rely on female sterilization, but a mere 13% do so as the average country. The high figure is due to India’s 67% of users on female sterilization, followed by China’s 34%, which together represent two-thirds (69%) of the region’s population. Table 1 shows that the difference is reversed for the injectable: it is not important in India and China, but it is very important in Indonesia, the region’s third largest country. The total rows give the overall contrasts for each of the 8 methods, including the large difference for female sterilization.

3. In sub-Saharan Africa, a Hormonal Method Progression From Oral Contraceptives to Injectables to Implants Is Evident

The sub-Saharan Africa region is especially relevant for contraceptive dynamics, as it shows the greatest amount of change as countries move toward modifications in the method mix.

Historically in sub-Saharan Africa, hormonal methods have dominated, apart from traditional methods. In the 1970s and 1980s, such use consisted largely of oral contraceptives. But with U.S. Food and Drug Administration approval of the

### Table 1. Method Mix for Latest Surveys to Compare Unweighted and Weighted Results

<table>
<thead>
<tr>
<th></th>
<th>Female Sterilization</th>
<th>Male Sterilization</th>
<th>IUD</th>
<th>Implant</th>
<th>Injectable</th>
<th>Pill</th>
<th>Condom</th>
<th>Traditional</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Asia</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unweighted</td>
<td>13.4</td>
<td>1.9</td>
<td>24.8</td>
<td>2.5</td>
<td>13.9</td>
<td>17.8</td>
<td>13.2</td>
<td>12.6</td>
<td>100.0</td>
</tr>
<tr>
<td>Weighted</td>
<td>38.7</td>
<td>2.6</td>
<td>22.0</td>
<td>1.0</td>
<td>6.6</td>
<td>10.0</td>
<td>10.7</td>
<td>8.4</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Latin America</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unweighted</td>
<td>29.1</td>
<td>1.2</td>
<td>8.4</td>
<td>1.9</td>
<td>16.4</td>
<td>19.1</td>
<td>14.3</td>
<td>9.5</td>
<td>100.0</td>
</tr>
<tr>
<td>Weighted</td>
<td>30.8</td>
<td>2.6</td>
<td>9.2</td>
<td>2.5</td>
<td>10.7</td>
<td>23.2</td>
<td>13.2</td>
<td>7.8</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Middle East/North Africa</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unweighted</td>
<td>7.2</td>
<td>0.3</td>
<td>20.2</td>
<td>0.2</td>
<td>4.6</td>
<td>29.4</td>
<td>8.6</td>
<td>29.5</td>
<td>100.0</td>
</tr>
<tr>
<td>Weighted</td>
<td>7.7</td>
<td>0.6</td>
<td>23.4</td>
<td>0.3</td>
<td>5.6</td>
<td>32.0</td>
<td>8.2</td>
<td>22.2</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Sub-Saharan Africa</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unweighted</td>
<td>3.9</td>
<td>0.2</td>
<td>3.4</td>
<td>11.3</td>
<td>30.8</td>
<td>21.4</td>
<td>12.0</td>
<td>16.9</td>
<td>100.0</td>
</tr>
<tr>
<td>Weighted</td>
<td>3.5</td>
<td>0.1</td>
<td>3.1</td>
<td>13.9</td>
<td>35.7</td>
<td>18.9</td>
<td>7.8</td>
<td>17.0</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unweighted</td>
<td>11.6</td>
<td>0.8</td>
<td>12.0</td>
<td>5.5</td>
<td>19.6</td>
<td>22.0</td>
<td>11.9</td>
<td>16.6</td>
<td>100.0</td>
</tr>
<tr>
<td>Weighted</td>
<td>29.1</td>
<td>2.0</td>
<td>17.5</td>
<td>3.3</td>
<td>12.0</td>
<td>14.8</td>
<td>10.2</td>
<td>11.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Abbreviation: IUD, intrauterine device.
injectable DMPA in 1994, injectables progressively became the predominant method in many countries. Implants first appeared with the approval of 6-capsule Norplant in 1990, followed by more advanced implants. Yet, provision of implants remained fairly modest, constrained both by fairly high cost and a limited service delivery infrastructure to provide them. However, price/volume guarantees negotiated between donors and the 2 major implant manufacturers in 2012 and 2013 reduced the price dramatically. Moreover, improved service delivery mechanisms, notably mobile service delivery and social franchising, vastly increased implant availability. The high and increasing prevalence of the implant (and its percentage of market share) is due not only to high adoption rates, but to the long continuation of use that the implant offers. However, after the recommended period of use, removals and reinsertions are needed, so a result of the growing numbers of users is that implant removals will accelerate, as noted by Christofield and Lacoste.

A good example of hormonal progression is Ghana. The leading method in the 1970s and 1980s was the oral contraceptive; it was overtaken by the injectable in the mid-2000s, which in turn was overtaken by the implant by 2017 (Figure 3). Currently, the shares are pill, 16%; injectable, 26%; and implant, 28%. The general hormonal progression pattern is evident in at least 21 other countries: Angola, Benin, Burkina Faso, Burundi, Chad, Ghana, Guinea, Guinea Bissau, Liberia, Malawi, Mali, Niger, Nigeria, Rwanda, Senegal, Sierra Leone, Timor Leste, Tanzania, Togo, Uganda, and Zambia.

The latest entry in hormonal method choice is subcutaneous injectable DMPA or DMPA-SC. It provides a lower dose of DMPA in an approach that is more conducive to community service delivery and even to self-injection. DMPA-SC is already becoming popular in several African countries due partly to the self-injection option.

### 4. Condom Use for Contraception Has Increased in Some Countries With High HIV Prevalence

Worldwide, HIV prevalence is highest in sub-Saharan African countries. Not surprisingly, with the advent of HIV, condom use has risen to substantial shares of all contraceptive use in some of those countries. For Botswana, Lesotho, and eSwatini (formerly Swaziland), condoms are the first or second most widely used contraceptive method; their shares of the method mix are 69%, 37%, and 28%, respectively. Several other countries have relatively high condom shares: Angola (23%), Namibia (22%), and South Africa (16%). In contrast, in other countries, the condom method share is only in the single digits: Zimbabwe (6%), Mozambique (6%), and Malawi (3%). It is likely that condom use is higher than these figures indicate, since some women are reluctant to admit condom use; also, when 2 modern methods are
reported including the condom, the rule is to classify such clients only under the other modern method.

### Method Mix and Skew According to Region and Country

#### 1. The Predominant Method Differs by Region

For all countries, as noted above, the most widely used method is female sterilization (29%), followed much lower by the IUD (18%). The pill (15%), injectable (12%), traditional methods (11%), and condom (10%) follow. The smallest percentages correspond to implants (3%) and vasectomy (2%) (Figure 4, total bars, and Table 1, weighted totals).

This overall perspective masks the remarkable fact that the leading methods differ considerably by region (weighted data) and country: female sterilization in Asia (39%) and in Latin America (31%), the pill in the Middle East/North Africa (32%), and the injectable in sub-Saharan Africa (36%). Within individual countries, the shares vary quite widely.

Why these sharp disparities? The share of each method reflects each region’s own balance of supply and demand influences over time. The sterilization share builds up gradually from annual adoptions over past years, during which those influences would have changed; the same is true for the other long-acting methods of the IUD and implant. On the other hand, current users of the resupply methods (condoms, pills, injectables) come largely from adoptions in the recent past since their average use time is relatively short; therefore, their use is more sensitive to recent influences, such as supply interruptions and shifting method preferences. Disparities in the family planning environment are large and fundamentally different in countries as dissimilar as India and Mali, and the result is a blend of cultural background, donor involvement, provider priorities, cost, access, and public response to the methods offered. In general, there is variety in pattern but consistency in a region over time.

#### 2. Method Skew Persists Over Time, but the Evenness of Method Mix Varies Greatly by Country

The number of countries with method skew has remain unchanged in recent years. Evidence from the most recent surveys shows that in these 113 countries, 34 countries (or 30%) show a skewed method mix, the same as the 30% found by Bertrand et al.\textsuperscript{26} and slightly lower than the 35% reported by Sullivan et al.\textsuperscript{27} In short, close to a third of developing countries still have a skewed method mix.

### Table 2. The 34 Countries That Have a Method Skew (>50%) as of the Most Recent Survey and Method Causing the Skew, Based on Women Married or in Union

<table>
<thead>
<tr>
<th>Method</th>
<th>Country</th>
<th>Skew (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Injectable</td>
<td>Ethiopia</td>
<td>64.4</td>
</tr>
<tr>
<td></td>
<td>Liberia</td>
<td>62.8</td>
</tr>
<tr>
<td></td>
<td>Haiti</td>
<td>61.7</td>
</tr>
<tr>
<td></td>
<td>Sierra Leone</td>
<td>54.3</td>
</tr>
<tr>
<td></td>
<td>Myanmar</td>
<td>52.9</td>
</tr>
<tr>
<td></td>
<td>Mozambique</td>
<td>51.9</td>
</tr>
<tr>
<td></td>
<td>Indonesia</td>
<td>51.8</td>
</tr>
<tr>
<td></td>
<td>Madagascar</td>
<td>51.1</td>
</tr>
<tr>
<td></td>
<td>Malawi</td>
<td>50.8</td>
</tr>
<tr>
<td>Traditional</td>
<td>Azerbaijan</td>
<td>76.8</td>
</tr>
<tr>
<td></td>
<td>South Sudan</td>
<td>65.7</td>
</tr>
<tr>
<td></td>
<td>DR Congo</td>
<td>64.8</td>
</tr>
<tr>
<td></td>
<td>Armenia</td>
<td>51.9</td>
</tr>
<tr>
<td></td>
<td>Libya</td>
<td>51.5</td>
</tr>
<tr>
<td></td>
<td>Bahrain</td>
<td>51.3</td>
</tr>
<tr>
<td></td>
<td>Mauritius</td>
<td>50.7</td>
</tr>
<tr>
<td>Pill</td>
<td>Sudan</td>
<td>77.6</td>
</tr>
<tr>
<td></td>
<td>Algeria</td>
<td>77.8</td>
</tr>
<tr>
<td></td>
<td>Morocco</td>
<td>74.7</td>
</tr>
<tr>
<td></td>
<td>Saudi Arabia</td>
<td>62.0</td>
</tr>
<tr>
<td></td>
<td>Zimbabwe</td>
<td>61.7</td>
</tr>
<tr>
<td></td>
<td>Mauritania</td>
<td>59.8</td>
</tr>
<tr>
<td></td>
<td>Laos</td>
<td>50.6</td>
</tr>
<tr>
<td></td>
<td>Turkmenistan</td>
<td>87.5</td>
</tr>
<tr>
<td></td>
<td>Uzbekistan</td>
<td>80.0</td>
</tr>
<tr>
<td></td>
<td>Tajikistan</td>
<td>64.4</td>
</tr>
<tr>
<td></td>
<td>Kyrgyzstan</td>
<td>55.6</td>
</tr>
<tr>
<td></td>
<td>Kazakhstan</td>
<td>54.4</td>
</tr>
<tr>
<td></td>
<td>Egypt</td>
<td>51.5</td>
</tr>
<tr>
<td>Female Sterilization</td>
<td>India</td>
<td>67.7</td>
</tr>
<tr>
<td></td>
<td>Dominican Rep.</td>
<td>58.6</td>
</tr>
<tr>
<td></td>
<td>El Salvador</td>
<td>51.7</td>
</tr>
<tr>
<td>Condom</td>
<td>Hong Kong</td>
<td>70.0</td>
</tr>
<tr>
<td></td>
<td>Botswana</td>
<td>69.3</td>
</tr>
</tbody>
</table>

The marked differences in method mix reflect each region’s own balance of supply and demand influences over time.
In the 34 countries with method skew, the leading method differs considerably. As shown in Table 2, the number of countries skewed toward each method is injectable (9), traditional methods (7), pill (7), IUD (6), female sterilization (3), and condom (2). Table 2 also shows the extent of method skew in each country. In no country does male sterilization or the implant have a share more than 50%, although the share for the implant has reached 46% in Burkina Faso. Also noteworthy, in half (17) of the countries, the method skew exceeds 60%.

Returning to the AD values as a measure of the evenness of the mix, we find that the 113 countries follow a bell-shaped curve, with a roughly normal distribution. Around the AD median of 11.8, about half of countries (65) are in a middle range, falling between ADs of 9.9 and 13.7, and 97 are within the wider range of ADs 8.6 to 15.0. A few are at relatively extreme values; for example, Nepal in the low range with an AD of only 6.6, and Egypt in the high range with an AD of 14.0. Those in the high range contain the especially skewed cases.

Total Contraceptive Prevalence Is Not Related to the Evenness of the Mix

Previous research has indicated that increasing the number of available methods results in higher contraceptive prevalence, but that can either increase or decrease the evenness of the mix. Based on the 113 most recent surveys, we found no statistically significant relationship (R²=0.0065, P=.95) between the evenness of method mix as measured by the AD and contraceptive prevalence (Figure 5). As the CPR rises, the AD values do not systematically change. There is a large variation in the AD values at any level of the CPR.

Several reasons appear to account for this lack of association. First, some countries, such as China or Vietnam, with high CPRs rely on only 1 or 2 modern methods, showing a highly skewed method mix. Other countries, such as Niger and the Democratic Republic of the Congo, are also highly skewed, but at low CPR levels. Additional countries at middle CPRs vary considerably in the spread of their methods, some with narrow spreads and others with wide ones. All this reflects regional disparities in method access and choice as well as other factors.

DISCUSSION

This analysis shows at least 3 positive trends: a decline in the shares held by traditional methods in favor of more effective contraceptives, a “hormonal progression” in sub-Saharan Africa with countries moving from pills to injectables and in many cases on to widespread implant use, and the increased use of condoms in some countries with

![Figure 5. Relationship Between the Measure of Average Deviation and Contraceptive Prevalence Rate, 113 Latest Surveys](image-url)
high HIV prevalence. Yet, challenges remain. Despite more than 5 decades of international family planning, traditional methods represent an average of 17% of the method mix in the 113 countries analyzed, or 11% of all users. And close to a third (30%) of countries still report method skew, with over half of all use by a single method.

Two unexpected findings are that a more even method mix is not associated with a higher CPR, and that the leading contraceptive methods differ considerably more among regions than we would have anticipated.

The current mix is a function of 2 dissimilar dynamics: use of the long-term methods is an accumulation of adoptions over past years, whereas use of short-term methods comes from recent starts, due to their shorter continuation rates. Therefore, the impact of current program initiatives and other determinants of use can be considerably greater among the short-term methods.

Some countries have implemented deliberate measures to diversify method mix. An intensive effort in parts of 5 crisis-affected countries (Chad, the Democratic Republic of the Congo, Djibouti, Mali, and Pakistan) to widen access to several methods resulted in 61% of clients selecting implants and IUDs. In Indonesia, community-led advocacy efforts implemented in the 6 Improving Contraceptive Method Mix project districts yielded increases in uptake of long-acting and permanent contraceptive methods, against a national context in which about half of users rely on the injectable. Yet, elsewhere such initiatives have failed to change the mix, such as efforts in Morocco in the 1990s to increase IUD use in a “pill” country. Despite efforts to encourage the uptake of vasectomy, its use has fallen sharply wherever it had claimed a significant share of use; currently the highest share is 10% in Nepal, 5% in Brazil and Colombia, and close to zero in many developing countries.

What explains the persistence of method skew in some countries? The 34 countries we found with skew are nearly the same as those in the 2006, 2014, and 2015 reviews. Method mix is like a slow-moving ship: it is possible to change direction only over time. It is often difficult to disentangle the 2 main categories of factors that influence skew: limitations on the supply side (lack of access to a wider range of contraceptives, beset by stock-outs, cost barriers, and provider biases) versus those on the demand side, including ingrained societal preferences. Is the high level of female sterilization in India or the Dominican Republic the result of constrained supply of alternative methods, normatively influenced demand, or both?

The above analyses allow us to better understand the current status of method mix, its evolution over time, and its diversity by region and country. Yet, key questions remain. First, to what extent is continuous method skew a problem in countries with high CPRs? Numerous countries have CPRs above 60% and are skewed by the 50% rule: Dominican Republic, Mexico, and El Salvador for female sterilization, Morocco and Zimbabwe for the pill, and near cases for the IUD: China and Vietnam with 48% of use on the IUD. We are unaware that any of these countries are taking action to improve the evenness of the mix.

Second, is it really a problem if a country moves toward greater method skew after the introduction of a new method, if the method enlarges choice and helps meet the needs of clients? For example, in Burkina Faso 46% of users now rely on implants, and other sub-Saharan African countries are moving in this direction.

Third, in the absence of an “optimal” or “ideal” method mix, are there measures that better capture the balance in contraceptive method mix that some program managers and donors seek and that are believed to better meet clients’ diverse needs? Bertrand et al. proposed using the real-life experience of countries that come closest to having a fully balanced method mix and also have at least a moderately high CPR, defined as 25%. Yet, in the absence of a widespread initiative to improve method mix, any method to improve the measurement of “balance” in method mix seems to lack programmatic relevance.

Another approach would be to examine possible relationships between family planning program effort measures and the characteristics of the mix. If strong programs best service the needs of clients, the resulting mix may be closer to a preferred standard. Such work would need to take into consideration the vast divergence among regions in predominant methods.

**Limitations**

Regarding limitations in this work, one relates to the surveys available. The number of surveys per country varied from 1 to 18, which decreased the sensitivity of the time trends in countries with few surveys. Also, the surveys were not conducted in the same years or at a constant interval, and we included multiple types of surveys (e.g., DHS, MICS, CPS, PMA2020) with their dissimilar methodologies. However, concerns about the latter were allayed by the reruns done with only the DHS surveys. Also, the surveys were not conducted in the same years or at a constant interval, and we included multiple types of surveys (e.g., DHS, MICS, CPS, PMA2020) with their dissimilar methodologies. However, concerns about the latter were allayed by the reruns done with only the DHS and MICS types, which gave very similar results to those produced by the full set.

Our primary focus on method mix resulted in less attention to prevalence. In countries where...
total prevalence is quite low, the mix among the 8 methods is less stable over time, and the share estimates are subject to greater sampling error. Total prevalence has risen in many countries, so that a method can lose share and still keep the same level of prevalence. Historically, countries have moved for example from a high share of traditional methods toward lower shares, even while the absolute level of their prevalence may have changed little. Wanting to focus especially on trends in the mix both overall and by regions, we gave less attention to the complexity of method mix as it occurs in particular countries. Nor did we analyze the relationship between method mix and economic status of countries by their GDP per capita or similar measures.

We did not undertake a separate analysis of method availability as a determinant of the mix, as beyond our scope. Measures of availability are found in the FP2020 annual report for the 69 poorest countries in the world and in a study of national family planning program efforts in more than 80 developing countries. Any analysis of the relationship of availability to other measures must contend with the problem that data are not always available for the year corresponding to the latest nationally representative survey; moreover, “availability” has several dimensions including geographic access, cost, and quality of care at the source of each method.

## CONCLUSION

A future step in researching method mix involves more in-depth analysis of the methods that produce the unevenness in method mix in relation to total contraceptive prevalence. Our analyses do not address the complex relationships among choice, total prevalence of use, and the various mix patterns. Most use in most countries is accounted for by 2 to 3 methods. Limited choices only partly account for that since consumer preferences enter in, and a full choice of many methods might not alter the prevailing pattern. Nevertheless, past experience confirms that the addition of more methods to a narrow mix increases prevalence, up to some limit. Further work into the history of which methods and at what prevalence levels would be of interest.

**Acknowledgments:** The authors wish to thank Rebecca E. Rosenberg for editorial assistance. John Ross acknowledges support from the Track20 Project of Avenir Health, which is supported by the Bill and Melinda Gates Foundation for his work on this article.

**Competing interests:** None declared.

## REFERENCES


Peer Reviewed

Received: May 20, 2020; Accepted: September 22, 2020; First published online: November 6, 2020


© Bertrand 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-20-00229
Health Care Worker Preferences and Perspectives on Doses per Container for 2 Lyophilized Vaccines in Senegal, Vietnam, and Zambia

Natasha Kanagat, Kirstin Krudwig, Karen A. Wilkins, Sydney Kaweme, Guissimon Phiri, Frances D. Mwansa, Mercy Mvundura, Joanie Robertson, Debra Kristensen, Abdoulaye Gueye, Sang D. Dao, Pham Q. Thai, Huyen T. Nguyen, Thang C. Tran

Key Findings

- Health care workers (HCWs) in all 3 countries preferred containers with fewer doses for reconstituted vaccines such as BCG and measles-containing vaccine.
- HCWs believed that containers with fewer doses of these vaccines could reduce wastage and missed vaccination opportunities.
- HCWs were more willing to open a vial for every eligible child when using containers with fewer doses.

Key Implications

- Policy makers should consider HCW perspectives when deliberating a change in policy on vial size since HCWs have to balance concerns about open vial wastage with the guidance to open a vial to vaccinate every eligible child.
- Program managers should consider shifting to containers with fewer doses for vaccines without preservatives to assuage HCW concerns about opening vials for every eligible child while managing wastage.

ABSTRACT

Introduction: Limited information exists on health care workers’ (HCWs) perceptions about use of multidose vaccine vials and their preferences about doses per container (DPC). We present findings from qualitative studies conducted in Senegal, Vietnam, and Zambia to explore HCWs’ behavior regarding opening vials and their perceptions and preferences for the number of doses in vials of BCG and measles-containing vaccine (MCV). Zambia and Senegal currently offer MCV in 10-dose vials and BCG in 20-dose vials; 10-dose vials are used for both vaccines in Vietnam. Unused doses in vials of these reconstituted vaccines must be discarded within 6 hours.

Methods: Key informant interviews (KII) were conducted with frontline HCWs in Senegal, Vietnam, and Zambia. In Senegal and Vietnam, the KII were conducted as part of broader formative research; in Zambia, KII were conducted in control districts using 10-dose MCV vials only and in intervention districts that switched from 10- to 5-dose vials during the study. During analysis, themes common to all 3 countries were synthesized. Critical themes relevant to country contexts were also examined.

Results: HCWs in all 3 countries preferred containers with fewer doses for BCG and MCV to reduce wastage and increase the likelihood of vaccinating every eligible child. HCWs in Senegal and HCWs using 10-dose vials in Zambia reported sending unvaccinated children away because not enough children were present to warrant opening a new vial. In Vietnam, where sessions are typically held monthly, and in Zambia when the 5-dose vials were used, almost all HCWs reported opening a vial of MCV for even 1 child.

Discussion: HCWs prefer vials with fewer DPC. Their concerns about balancing coverage and wastage influence their decisions to vaccinate every eligible child; and their perspectives are crucial to ensuring that all target populations are reached with vaccines in a timely manner.

INTRODUCTION

Many vaccines administered in low- and middle-income countries are purchased in multidose vials (MDVs) and can contain between 2 and 20 doses per vial.
or container.\textsuperscript{1,2} Several countries buy vaccines in MDVs because compared with single-use vials, MDVs sell at a lower price per dose; require lower cold chain, storage, and transport capacity; and generate less waste.\textsuperscript{3} Some vaccines contain preservatives, whereas others do not. Under the World Health Organization’s (WHO’s) multidose vial policy,\textsuperscript{4} remaining doses in open vials of vaccines with preservatives can be used for up to 28 days after opening, as long as storage and proper handling conditions are met. However, vaccines without preservatives must be used in a much shorter time frame. Vaccines such as BCG, measles-containing vaccine (MCV), and yellow fever vaccines do not contain preservatives, and they must be discarded within 6 hours of reconstitution or at the end of a session, whichever comes first. Health care workers (HCWs) in low- and middle-income countries who administer these vaccines to their target populations are therefore responsible for deciding when to open a vial, knowing that if not all doses are used within a short frame of time, they will have to be discarded, resulting in open-vial wastage.

This article focuses on vaccines without preservatives. HCWs have to balance the expectation that they will vaccinate every child with the concerns about open-vial wastage. Open-vial wastage tends to increase with vaccines that have more doses per container (DPC) when the immunization session sizes are small.\textsuperscript{5} Limited information exists on HCWs’ opinions about the desired DPC and how DPC informs their decisions on when to vaccinate. Studies by Wallace et al.\textsuperscript{6} and Hutchins et al.\textsuperscript{7} suggest that HCWs’ hesitancy to open a multidose vaccine vial to avoid vaccine wastage contributed to missed opportunities for vaccination (MOVs). HCWs’ behavior regarding opening vials is critical to addressing MOVs, which emphasizes efforts to reach eligible children at all immunization sessions, including outreach, to identify and reduce opportunities missed at the health facility level on a day-to-day basis. MOVs can result in inadequate protection against disease.

The qualitative studies described here (see Methods section) were part of a larger multicountry project to improve the evidence base on HCWs’ decision making relative to DPC. This article focuses on qualitative findings on HCWs’ perspectives on BCG and MCV, obtained through formative research in Senegal and Vietnam and a prospective study in Zambia. The project also conducted household surveys to examine immunization coverage, administered facility surveys to conduct cost-effectiveness analyses, and studied routine administrative data from facilities to assess vaccine wastage. The findings from those studies are forthcoming in other journals.

In this article, we report on the BCG and MCV vaccines since they are common to all 3 countries, are supplied in MDVs, and must be discarded after 6 hours. Zambia and Senegal use a 10-dose measles-rubella (MR) vial and a 20-dose BCG vial. Vietnam uses 10-dose MCV (both measles and MR) and BCG vials; measles first dose is given at 9 months and MR vaccine (second dose) is given at 18 months.

In Senegal, the MR first dose is given at 9 months and is coadministered with yellow fever vaccine. The MR second dose is given at 15 months. All routine childhood immunizations are given during fixed and outreach sessions. Health facilities conduct fixed sessions that are held anywhere from daily to monthly, depending on the catchment population and size of the facility. Outreach sessions vary in frequency, depending on the number of outreach locations, availability of staff to conduct outreach, and other factors. In Vietnam, fixed sessions are held once or twice a month in most of the country, although in some districts immunization is organized once weekly. Outreach sessions are not conducted everywhere, and they vary in frequency where offered. In Zambia, MR is given at 9 months and 18 months of age. Routine childhood immunizations in Zambia are given during fixed and outreach sessions. Health facilities hold sessions anywhere from daily to monthly depending on the catchment population, size of the facility, availability of staff to conduct outreach, and other factors.

In Zambia, the guidance is to open a vial for every eligible child, and WHO’s multidose vial policy is followed. Vaccines with preservatives can be kept for up to 28 days, while vaccines without preservatives (i.e., BCG, MCV, and yellow fever) must be discarded 6 hours after reconstitution or at the end of an immunization session, whichever comes sooner.

In Senegal, the national level does not give guidance to health facility staff on how many eligible children must be present at a session before HCWs can open a vial. As in Zambia, WHO’s multidose vial policy is followed.

In Vietnam, sessions are held monthly in most health facilities; therefore, the policy for all vaccines is to discard all remaining doses in opened vials at the end of each session.
**METHODS**

The qualitative study in Senegal was conducted as part of a broader formative research study (Box). Sixty health facilities (HFs) were included in the study and 1 HCW per health facility participated in qualitative interviews. Health facilities were selected through stratified random sampling based on rural or urban locations, size of birth cohort served, and distance from the district vaccine store. The study was conducted in 2 regions, Louga and Ziguinchor, selected because their vaccine coverage rates are below national coverage rates.

The qualitative study in Vietnam was conducted as part of a broader formative research study that included 30 health facilities. Thirty HCWs (1 per HF) participated in the qualitative interviews. The study was conducted in 4 provinces in 2 regions in Vietnam: Dien Bien, Tuyen Quang, and Yen Bai in Northern Region, and Dak Lak in the Central Highlands Region. Dak Lak Province was chosen specifically because the immunization session is organized once every week, whereas in the other sites the immunization sessions are conducted once or twice a month. The researchers wanted to assess whether HCWs’ perspectives differed across sites with different session frequencies. In Northern Region, 2 districts in each province and 3 communes in each district were selected. In Central Highlands Region, the researchers focused on the one district that was conducting weekly immunization, and they included 12 communes from that district and therefore oversampled in the district with weekly immunization. Thirty HFs were purposively sampled, taking into account coverage and different service delivery models. In the Northern Region, health facilities were selected either because they had coverage rates below 90% or the lowest coverage in the district. In some cases, even if they had the lowest coverage rates in the district, their coverage was over 90%. Most HFs in the Northern Region are in rural settings, whereas the facilities in the Central Highlands Region contained more of a mix of rural, peri-urban, and urban sites. In Vietnam, HCWs were asked about measles and MR separately, because their vaccine schedule requires measles for first dose and MR for second dose.

The qualitative study in Zambia was conducted as part of a broader implementation research study. In Zambia, 90 HCW interviews were conducted across 14 districts in 2 provinces. For the implementation research, all districts were paired according to average population size per HF and the number of HFs. From each pair, 1 district was randomly assigned to the intervention, while the other district served as the control. During the implementation period, all HFs in the intervention district received 5-dose vials of MCV, while the

---

**BOX. Introduction to the Dose Per Container Partnership Project Under Which the Studies Presented in the Article Were Conducted**

The global effort to protect all people from vaccine-preventable diseases has historically leveraged multidose containers in low- and middle-income countries to offer lower prices and reduce the constraints on cold chain space. However, as newer, more expensive vaccines are introduced in multidose formats, the burden of cost efficiency potentially moves from the national-level to the health care worker.

To achieve maximum utilization of every dose in a vial and depending on the country’s policies, health care workers need to be strategic about when to open a vial and be diligent about how they care for open vials. In addition, they have to be more active with community outreach and communication to ensure optimal attendance and timely immunization. For these reasons, the number of doses per container (DPC) can hinder a country’s ability to achieve timely and equitable coverage including reaching the urban poor or rural remote. DPC can also influence additional factors like vaccine safety, system costs, supply chain, and wastage.

The Dose Per Container Partnership (DPCP) was a multicountry project that aimed to support vaccine product and program decision making to include considerations of DPC to optimize equitable, timely, safe, and cost-effective coverage.

The Partnership implemented country-level research in several countries, including Senegal, Zambia and Vietnam, to generate new evidence on the impact of DPC decisions on an immunization system, to explore current decision making on DPC options, and to inform country and global decisions on vaccine procurement. The Partnership has produced case studies on decision making and multicountry research, as well as videos, resource guides, and briefs on various aspects of DPC.

DPCP is funded by the Bill & Melinda Gates Foundation, led by JSI Research & Training Institute, Inc. and jointly implemented in partnership with PATH, the Clinton Health Access Initiative, the Highly Extensible Resource for Modeling Event-Driven Supply Chains modeling team, and the International Vaccine Access Center.
control group continued to use the 10-dose vials. The HFs receiving the 5-dose vials were oriented to the new vial size, but no other technical support was offered over the 1-year implementation period, to minimize influencing the HCWs’ behavior. The project wanted to keep the settings as neutral as possible to allow us to observe differences in behavior at endline. HCW interviews for the qualitative study were conducted at baseline (32), midline (16), and endline (42).

Interview teams verified that respondents from all 3 countries were responsible for providing routine immunizations to ensure they had an understanding of immunization service delivery. HFs were selected to ensure representation across large, medium-sized, and small facilities in urban and rural locations within each district.

At baseline, none of the districts in Zambia had switched over to the 5-dose vials, so all 32 interviews were conducted with HCWs who were using the 10-dose vials. The midline and endline interviews, however, were conducted after the districts were divided into those receiving 10-dose vials and those receiving 5-dose vials, so midline and endline interviews were held with HCWs in districts receiving 5-dose vials to document their experiences using the new presentation.

In all 3 countries, contracted local data collectors gathered data with oversight and supervision by the organization leading the country study. Data analysis for each country was done separately. For Senegal and Vietnam, responses to the qualitative surveys were analyzed in Excel. For the questions with predefined response options, the responses were counted based on the response options. For the open-ended questions, the key themes from the responses were also tabulated and reported. For Zambia, all transcripts were uploaded into NVivo 11, a qualitative data management software. The qualitative team generated an initial set of codes derived from the research questions to analyze the data. All codes were accompanied by code definitions. The initial set of codes comprised major thematic categories, which were refined through analysis, and subcategories (i.e., subcodes) were developed through iterative analysis. For this article, all the country reports and briefs generated from the separate analysis were then reviewed to identify major themes common to all 3 countries and summarize findings. We also highlight country-specific findings.

The formative research studies in Senegal and Vietnam were determined by PATH’s Research Determination Committee not to be human subjects research. Approvals were obtained from the Ministry of Health/Senegal Ethics Committee and the Ethics Committee of the Vietnamese National Institute of Hygiene and Epidemiology. The implementation study in Zambia received ethical approval from the Biomedical Research Ethics Committee of the University of Zambia.

### RESULTS

#### HCW Perceptions on Reducing Missed Opportunities When Using 5-Dose Vials

All HCWs were asked (1) whether they vaccinated every eligible child each time the child was at the health facility; (2) if MCV and BCG were offered at every session; and (3) if they opened a vial of these vaccines at a session irrespective of the number of eligible children present. These questions were asked to assess whether concerns about opening a vial for only 1 child or a few children resulted in HCWs either not opening the vial or waiting for a minimum number of children being present to justify opening it.

In Senegal, MCV, BCG, and yellow fever vaccines were not offered at every immunization session. Most respondents said that at least 5 children had to be present for MCV and yellow fever vaccine and 10 children had to be present for BCG before they would open a vial. When fewer children attended a session, HCWs asked them to return on another date when the next session was scheduled. In Senegal, the majority of HCWs recalled turning away a caregiver and child from a vaccination session at least once in the last 3 months. Similar findings were also observed for yellow fever vaccine and MR, which is also in 10-dose vials and is coadministered with MR.

*I programmed them for the next session, it’s for tomorrow. I have recorded their coordinates [location] and the relays [community health workers] take care to find them, and if they do not come, we call them on the telephone.—Senegal*

In Vietnam, the majority of facilities conduct sessions once a month, and all vaccines are offered at each session. Therefore, the majority of HCWs reported that they opened a vial for every child and were willing to open a vial during sessions when only 1 child was eligible, regardless of potential wastage. Most HCWs in Vietnam did not recall sending children away during an immunization session because not enough children were present to warrant opening a vaccine vial. The few HCWs who
recalled sending children away mentioned that they advised the caregivers to bring the child back for the next session.

Sending children away without vaccination takes time because the caregivers have to bring their children back for the other immunization session.—Vietnam

In Zambia, over half of HCWs using 10-dose vials of MCV reported waiting for a minimum of 5 children before opening a vial, and a minimum of 10 children to open a vial of BCG. However, when HCWs using 5-dose vials were asked about their practices, they replied that they were less concerned about MCV wastage and felt more comfortable opening vials to vaccinate children. Most HCWs in Zambia using the 5-dose vial stated that they opened a vial each time an eligible child presented during a session and did not wait for a minimum number of children.

We can open the vial even when we have two children, we only lose three doses as compared to the time we were using ten-dose vials, this would make us lose eight doses.—Zambia 5-dose vial district

In districts that continued to use the 10-dose MCV vial in Zambia, the majority of HCWs recalled turning away a child from a vaccination session at least once in the past 3 months. In the Zambian districts using 5-dose vials of MCV, very few HCWs reported turning children away. Neither group of respondents had a system to track whether the children turned away were brought back to the facilities for vaccinations in the future.

Yes, because everyone is concerned on reducing the vaccine wastage. It is a reason why mothers are sent back and asked to come a different day when there are enough children to open the vial. This is so because everyone wants to reduce the wastage.—Zambia 10-dose vial district

### Balancing Coverage and Wastage

All HCWs were asked whether their supervisors considered coverage rates or wastage rates more important, since the supervisors’ belief would influence what the HCWs placed more emphasis on. The belief also could affect HCWs’ behavior if they offered certain vaccines at specific times to ensure adequate numbers of children to minimize wastage. Most HCWs from Senegal, Vietnam, and Zambia stated that their supervisors deemed coverage more important than wastage. However, if wastage rates were deemed higher than expected, HCWs reported that supervisors offered suggestions and strategies to mitigate wastage.

In Senegal, most HCWs knew the target wastage rates for each vaccine. Knowing the target affected which vaccines were offered at each session. To minimize wastage rates, HCWs reported not offering vaccines like BCG and MCV at each session. When asked how they ensured that a certain number of children were present for the BCG and MCV sessions, they replied:

Collect 10 children for BCG, MR [measles rubella], before opening the bottle [vial], using relays and badjene ngokh [community health workers] who will bring the children and remind parents of the RV [vaccination session].—Senegal

Most HCWs in Senegal and the 10-dose districts of Zambia also reported that MCV and BCG vaccines were not offered at every session. The reason for not offering these vaccines was to increase session sizes for these specific vaccines to reduce wastage.

In Vietnam, the session frequency determined whether all vaccines were offered at every session. At sites where immunization was offered once or twice a month, all vaccines were offered every time. At sites where immunization was offered more frequently, not all vaccines were offered every time, to avoid vaccine wastage.

BCG, measles, and Japanese encephalitis (JE) vaccines are injected once every 2 weeks. DPT and MR vaccines are injected once per month. It is because the number of children who need these vaccines is less than that of other vaccines.—Vietnam

By contrast, in the Zambian districts using 5-dose vials of MCV, the majority of HCWs reported offering MCV at every fixed session regardless of the number of children. The HCWs in the 5-dose districts did not know their wastage rates, but they believed that wastage had diminished with their use of the 5-dose vials, and they were therefore less concerned about opening the MCV 5-dose vial for fewer children. In all 3 countries, BCG was given on specific days, such as at postnatal sessions at health facilities or hospitals, or on a designated day per month, to ensure that a large number of children would be present and wastage could be limited.

Yes there are days when MCV and BCG is not given during outreach, for example you find two babies who have been delivered. Are you going to open that vial for BCG
just for those two? So in such cases we explain to them that we can’t offer them BCG or MCV at that particular moment. We then advise them to come to the center, especially on the last Thursday of the month.—Zambia 10-dose district

What Are HCWs’ Preferences for DPC?

All HCWs expressed a preference for a different vial size of BCG and MCV with fewer DPC. In Senegal, most HCWs preferred fewer DPC for these vaccines to reduce wastage, and many said that this could help to address the challenge of needing enough eligible children to warrant opening a vial, prevent dropouts, and provide services to hard-to-reach children. A couple of HCWs felt that fewer DPC might pose challenges for storage or transportation of vials. In Zambia, HCWs from the 5-dose districts also expressed a preference for fewer DPC vials for MCV; none of them wanted to return to using 10-dose vials. The majority of HCWs in the 5-dose districts preferred 5-dose vials, and the rest preferred fewer than 5 doses. In Vietnam, for BCG and MCV (currently in 10-dose vials), the majority of HCWs expressed a preference for a 1-dose vial, followed by a 5-dose and a 2-dose vial (Table).

Because of the mobile population here we often lose sight of children—having single dose presentations would permit us to vaccinate each child who presents.—Senegal

If vaccine was packed in single dose per vial, we could conduct vaccination in more days per month instead of doing in 1 day.—Vietnam health facilities conducting weekly sessions

It has made things easier for us in that you do not have to worry about babies not being immunized; it’s rare that we miss out any child. It has made our work easier; our minds are free that we are doing our job [immunizing] unlike the BCG.—Zambia 5-dose vial district

DISCUSSION

This multicountry study demonstrates that waiting for a minimum number of children before opening a vial of BCG and MCV could result in MOVs. Eliminating or greatly reducing MOVs is critical to achieving the Global Vaccine Action Plan 2011-2020 goal of “90% national coverage and 80% in every district or equivalent administrative unit, for all vaccines in the national immunization schedule.”

HCWs serve as critical intermediaries between communities and the health system, and they regularly make decisions about the management and delivery of vaccine services to achieve recommended coverage levels. Although many factors contribute to MOVs, it is vital to recognize the role of vials with fewer DPC for reconstituted vaccines in reducing instances of MOVs. As is seen in Vietnam, with the variation of timing and vaccines offered during immunization sessions, high-level structural decisions are made to balance coverage and wastage. This study also highlights the importance of providing HCWs with options that do not require sacrificing vaccination coverage or having high wastage. Current practices create a tension between expectations and ground realities, obligating HCWs to offer life-saving vaccines infrequently, turn away children, or risk not meeting expectations on wastage. This finding supports the conclusions of Wallace et al. that HCWs either take active measures to reduce wastage or feel some conflict when wastage is high. Separate quantitative analyses from this project confirm HCWs’ perceptions that fewer DPC will likely increase coverage children and reduce wastage. In Zambia, facilities using 5-dose vials had 47% lower wastage rates compared with those using 10-dose vials. An increase in coverage of MR first and second dose respectively by 5% and 3.5% in the districts using 5-dose vials was attributable to the intervention (i.e., the use of the 5-dose vials).

HCWs from all countries also reported turning away children if not enough children were present to warrant opening a vial, and in many cases, no system was in place to ensure that these children would be vaccinated later. This practice goes against WHO recommendations that vaccination programs include daily opportunities for vaccination with all vaccines, offering vaccination at every contact, including screening at curative consultations, even if there is only 1 child. This behavior also represents a MOV, requiring additional effort by caregivers and HCWs to follow up and increasing the chance that the child will not receive BCG or MCV. Similar findings have been documented in other low- and middle-income countries where BCG and MCV are offered less frequently than vaccines that do not have to be discarded after 6 hours of reconstitution, as a way to increase the number of children present per session before opening a vial.

Our findings are pertinent to current discussions on session sizes during the coronavirus disease (COVID-19) pandemic. Due to the pandemic, WHO guidance recommends frequent routine
immunization sessions of smaller size to reduce the risk of spreading the virus.14

As more countries consider changing their DPC for different vaccines, decisions should take HCWs’ perspectives into account. This approach is not always the norm. Other DPCP case studies on decision making on DPC in Ghana, Benin, Côte d’Ivoire, and the Democratic Republic of Congo showed that HCWs’ perspectives were notably absent.15–17

We recommend that future research continue to explore the causal links between HCWs’ practices related to vaccine wastage and their impact on vaccination coverage, MOVs, and cost implications. We also recommend additional research on HCWs’ preferences in other countries and settings to expand the body of evidence regarding HCWs’ decision making about opening vials.

**Limitations**
This study had limitations, including differences in study design between countries and different criteria used to select health facilities and key informants. Data analysis was also done by separate teams. In 1 country, respondents were purposefully selected, which may limit the generalizability of the results. However, the large sample used for the qualitative interviews in all 3 countries ensured that we got an appropriate and adequate number of respondents whose views likely represent those of the larger population of HCWs. In Zambia, we collected data at 3 different times to enable us to document behavior change in HCWs, especially in the districts that switched to using the 5-dose vials. The researchers tried to address these differences by ensuring that data on priority themes were collected across all countries, and that research protocols, data collection tools, and draft reports were shared among teams to establish a level of consistency in the data being collected and analyzed. Another limitation is that this study focused on relatively low-performing districts. However, our findings are likely also applicable to high-performing districts given that public-sector resources are always limited and that striking a balance between

---

**TABLE. Summary Findings on Vaccine Doses per Container and Health Care Workers’ Perceptions and Practices In 3 Countries**

<table>
<thead>
<tr>
<th>Theme</th>
<th>Summary Findings</th>
</tr>
</thead>
</table>
| HCW perceptions on reducing missed opportunities when using 5-dose vials | • Senegal: MCV, BCG, and yellow fever vaccines were not offered every time immunization sessions were held. During immunization sessions, HCWs reported that they waited for a minimum number of children before opening these vaccines. HCWs recalled turning away a caregiver and child at least once in the past 3 months.  
• Vietnam: Due to Vietnam’s session schedules, which are mostly once a month, most HCWs did not wait for a minimum number of children before opening vials. They did not recall turning away children in the past 3 months.  
• Zambia: In the districts using the 10-dose vials, HCWs waited for a minimum of 5 children to open an MCV vial and 10 children to open a 20-dose BCG vial. In the districts using the 5-dose vials, HCWs opened a vial each time an eligible child presented and did not report turning away a child. |
| Balancing coverage and wastage            | • In all 3 countries, although coverage was considered more important, HCWs reported that wastage was tracked very closely and they knew they had to minimize wastage as much as possible.  
• In Zambia and Senegal, HCWs did not offer MCV, BCG, or yellow fever vaccines (Senegal only) at every session due to concerns about wastage. The intent was to increase session sizes for these specific vaccines as a way of reducing open vial wastage.  
• In the facilities offering 5-dose vials in Zambia, HCWs believed that their wastage was lower, and they expressed less concern about opening the vial for fewer children compared with the facilities using the 10-dose vials. |
| HCWs’ preferences for DPC                 | • All HCWs expressed a preference for fewer DPC for BCG and MCV (and yellow fever for Senegal) to allow them to vaccinate eligible children, prevent dropouts, and not worry about wastage.  
• No HCWs in the districts in Zambia that used the 5-dose vials during implementation wanted to return to using the 10-dose vials. |

Abbreviations: BCG, bacille Calmette-Guérin; DPC, doses per container; MCV, measles-containing vaccine; HCW, health care worker.

---

As more countries consider changing their DPC for different vaccines, decisions should take HCWs’ perspectives into account.
vaccinating every child and limiting wastage will be a difficult decision for HCWs in both high- and low-performing districts.

**CONCLUSION**

This 3-country study contributed evidence on HCWs’ perceptions and preferences with regard to various DPC options for reconstituted vaccines. The results suggest that when balancing the mandate to achieve high coverage and reduce vaccine wastage, HCWs have to decide when to open a vial with more DPC. In all 3 countries, high coverage rates were considered more important than not with more DPC. In all 3 countries, high coverage rates were considered more important than not with more DPC. In all 3 countries, high coverage rates were considered more important than not with more DPC. In all 3 countries, high coverage rates were considered more important than not with more DPC. In all 3 countries, high coverage rates were considered more important than not with more DPC. In all 3 countries, high coverage rates were considered more important than not with more DPC.

**Acknowledgments:** We wish to acknowledge the Ministries of Health of Senegal, Vietnam, and Zambia, particularly the Expanded Programme on Immunization staff and health care workers who supported and participated in this research. We also thank the data collection teams who conducted interviews in each country. We thank the Bill & Melinda Gates Foundation for funding this important work and to the DPCP partner organizations and Technical Advisory Group for reviewing protocols and supporting the synthesis of results.

**Competing interests:** None declared.

**REFERENCES**


Préférences des professionnels de la santé et perspectives sur les doses par flacon de deux vaccins lyophilisés au Sénégal, au Vietnam et en Zambie

Messages clés

Lorsqu’ils fournissent des services de vaccination, les agents de santé trouvent un équilibre entre la nécessité d’atteindre une couverture élevée et celle de limiter le gaspillage de vaccins. Les travailleurs de 3 pays ont déclaré que des récipients contenant moins de doses de vaccin contre la rougeole et le BCG leur permettraient de vacciner tous les enfants qui se présentent, tout en réduisant les inquiétudes quant au gaspillage de vaccin.

Résumé

Introduction: Il existe peu d’informations sur les perceptions des agents de santé (AS) concernant l’utilisation des flacons de vaccins multidoses et leurs préférences en matière de doses par flacon (DPF). Nous présentons les résultats d’études qualitatives menées au Sénégal, au Vietnam et en Zambie pour étudier le comportement des agents de santé concernant l’ouverture des flacons et leurs perceptions et préférences quant au nombre de doses dans les flacons de BCG et de vaccin contre la rougeole. La Zambie et le Sénégal utilisent actuellement le vaccin contre la rougeole en flacons de 10 doses et le BCG en flacons de 20 doses; des flacons de 10 doses sont utilisés pour les deux vaccins au Vietnam. Les doses inutilisées des flacons de ces vaccins reconstitués doivent être jetées dans les 6 heures.

Méthodes: Des entretiens avec des informateurs clés ont été menés avec des agents de santé de première ligne au Sénégal, au Vietnam et en Zambie. Au Sénégal et au Vietnam, les entretiens ont été menées dans le cadre d’une recherche formative plus large; en Zambie, les entretiens ont été menées dans les districts de contrôle en utilisant uniquement des flacons de 10 doses de vaccin contre la rougeole et dans les districts d’intervention qui sont passés de flacons de 10 à 5 doses au cours de l’étude. Au cours de l’analyse, les sujets communs aux trois pays ont été synthétisés. Les sujets critiques pertinents pour les contextes nationaux ont également été examinés.

Résultats: Les travailleurs de la santé des trois pays ont préféré des récipients contenant moins de doses de BCG et de vaccin contre la rougeole afin de réduire le gaspillage et d’augmenter la probabilité de vacciner chaque enfant éligible. Les agents de santé au Sénégal et les agents de santé utilisant des flacons de 10 doses en Zambie ont déclaré avoir renvoyé des enfants non vaccinés parce qu’il n’y avait pas assez d’enfants présents pour justifier l’ouverture d’un nouveau flacon. Au Vietnam, où les séances ont généralement lieu tous les mois, et en Zambie, où les flacons de 5 doses ont été utilisés, presque tous les travailleurs de la santé ont déclaré avoir ouvert un flacon de Rougeole, même pour un seul enfant.

Discussion: Les agents de santé préfèrent les flacons contenant moins de dose par flacon. Leur souci d’équilibrer la couverture et le gaspillage influence leurs décisions de vacciner chaque enfant éligible; et leur point de vue est crucial pour garantir que toutes les populations cibles soient vaccinées en temps voulu.
Remote Mentorship Using Video Conferencing as an Effective Tool to Strengthen Laboratory Quality Management in Clinical Laboratories: Lessons From Cambodia

Grant Donovan, a Siew Kim Ong, b Sophanna Song, b Nayah Ndefru, c Chhayheng Leang, b Sophat Sek, b Patricia Sadate-Ngatchou, a Lucy A. Perrone c

Key Findings

- Utilization of a mixed-methods intervention design combining remote and in-person training, accompanied by close mentorship, contributed to successful implementation of quality management systems in participating laboratories.
- Laboratory participation time in video conference training activities correlated with better quality management system management and improved conformity to the ISO 15189 standard for medical laboratories.

Key Implications

- When identifying budget and policy priorities for health, policy makers should consider the beneficial impact of a sustained human resources training and mentorship program on laboratory quality improvement and service delivery efforts.
- Policy makers should particularly consider the potential efficiency and effectiveness of remote-access telementoring and teleconferencing to support online communities of practice for laboratory professionals because improved connectivity and knowledge sharing between professionals are essential for quality service delivery in a laboratory system.

ABSTRACT

Background: Providing professional development opportunities to staff working in clinical laboratories undergoing quality improvement programs can be challenged by limited funding, particularly in resource-limited countries such as Cambodia. Using innovative approaches such as video conferencing can connect mentors with practitioners regardless of location. This study describes and evaluates the methods, outputs, and outcomes of a quality improvement program implemented in 12 public hospital laboratories in Cambodia between January 2018 and April 2019. The program used mixed intervention methods including both in-person and remote-access training and mentorship.

Methods: Training outputs were quantified from the activity reports of program trainers and mentors. Program outcomes were measured by pre- and postimplementation audits of laboratory quality management system conformity to international standards. Variations in improved outcomes were assessed in relation to the time spent by laboratory personnel in video conference training and mentoring activity. An additional cross-sectional comparison described the difference in final audit scores between participating and nonparticipating laboratories.

Results: Laboratories significantly improved their audit scores over the project period, showing significant improvement in all sections of the ISO 15189 standard. Pre- and postaudit score differences and laboratory personnel participation time in remote mentoring activities showed a strong monotonic relationship. Average input per laboratory was 6,027 ± 2,454 minutes of participation in video conference activities with mentors. Audit scores of participating laboratories were significantly higher than those of laboratories with no quality improvement program.

Conclusion: Laboratories improved significantly in ISO 15189 conformity following structured laboratory quality management systems training supported by remote and on-site mentoring. The correlation of laboratory participation in video conference activities highlights the utility of remote video conferencing technology to strengthen laboratories in resource-limited settings and to build communities of practice to address quality improvement issues in health care. These findings are particularly relevant in light of the COVID-19 pandemic.

INTRODUCTION

Development of strong laboratory quality management systems (LQMS) is a key component of strengthening health systems for improved health care. Providing professional development opportunities to staff working in clinical laboratories undergoing quality improvement programs can be challenging due to limited funding, particularly in resource-limited countries such as Cambodia. Using innovative approaches such as video conferencing can connect mentors with practitioners regardless of location. This study describes and evaluates the methods, outputs, and outcomes of a quality improvement program implemented in 12 public hospital laboratories in Cambodia between January 2018 and April 2019. The program used mixed intervention methods including both in-person and remote-access training and mentorship.
outcomes and disease surveillance in resource-limited countries, and it requires standardization and strategic planning.\textsuperscript{1-3} ISO 15189 accreditation, which is the international standard for medical laboratory quality, provides standardization of LQMS requirements with a strong technical foundation for health, safety, and conformity.\textsuperscript{4,5} These standards are stringent, however, and have required a variety of approaches for laboratories with different resource availability and levels of development to achieve them.\textsuperscript{5,6} In Cambodia, a national effort to meet International Health Regulations and improve health services has culminated in an expansive national laboratory system to meet the diagnostic and surveillance needs of the country at both the national and provincial levels.\textsuperscript{7} The country has adopted national standards for medical laboratories, integrated a system of external quality assessments through private and public partnerships, and developed a national laboratory information system to improve surveillance and care, but structured quality improvement (QI) programs are limited to only a subset of laboratories. Expansion of these quality management training programs to meet international standards of quality was recommended in a series of laboratory assessments carried out between 2013 and 2016.\textsuperscript{7,8} One of these assessments measured 11 indicators of laboratory capacity, identifying a low average score of 36% in 22 laboratories, with indicators of LQMS averaging only 47% due to a lack of quality management systems, trained quality assurance managers, or continuous improvement practices.\textsuperscript{8} The implementation of structured, stepwise programs to improve quality management systems in national and provincial laboratories has been integral to improving laboratory quality and capacity in Cambodia.\textsuperscript{8,7} In 2001, the U.S. Centers for Disease Control and Prevention partnered with the Cambodia Ministry of Health (MOH) to implement the Strengthening Laboratory Management Toward Accreditation program using the Stepwise Laboratory Improvement Process Towards Accreditation (SLIPTA) audit tool, supported by quality management training and mentorship by trained QI professionals (BMLS Cambodia, unpublished presentation, 2018).\textsuperscript{9} International Training and Education Center for Health (I-TECH’s) QI program began in 2014 and was intended to expand access to LQMS training and implementation coaching nationally, delivered through a complementary package of training, mentorship, and technical assistance to MOH for national laboratory policy and guideline development. Success of these programs in Cambodia and globally has shown the impact that structured and mentored LQMS programs can have in resource-limited health care systems, holding promise for other such programs in Cambodia in the future.\textsuperscript{8,9} However, delivery of professional training and close mentorship for laboratories undergoing QI programs remains challenged by geographical and economic constraints. These challenges have prompted the use of modern video conferencing technologies to expand access to consultation from quality management professionals to distal facilities. Studies have shown that the use of these technologies, collectively known as telementoring, is an efficient and cost-effective tool to provide seasoned or specialized expertise to health professionals in remote or resource-limited medical facilities and has an impact on professional behavior and knowledge, as well as health outcomes.\textsuperscript{10-12} One recent program in Southeast Europe used monthly mentorship through telecommunication to improve laboratory quality in 5 countries and demonstrated measurable progress within the 6 laboratories supported.\textsuperscript{13} One of the most successful models of telementoring, Project ECHO (Extension for Community Healthcare Outcomes), has recently been expanded to laboratory strengthening, and the institute is now partnering with at least 5 major professional laboratory institutions to provide laboratory training and mentorship communities of practice globally.\textsuperscript{14} Research demonstrating the effectiveness of this model of remote training and mentoring for laboratory strengthening is limited, however, prompting a need for quantitative research.

**Program Description**

During the initial phase (Phase 1) of this project in Cambodia in 2014, a group of 12 participating laboratories received training and mentored technical support to implement an LQMS according to the newly published World Health Organization (WHO) Laboratory Quality Stepwise Implementation (LQSI) tool.\textsuperscript{15} Evaluation of that phase of the program indicated that consistent on-site mentoring in the local language with a stepwise action plan enhanced staff knowledge of LQMS implementation towards meeting the ISO 15189 standard, without interrupting regular laboratory services.\textsuperscript{8} The successful results of this training and mentoring approach led to an initiative by MOH to expand laboratory access to LQMS training and mentorship that prioritized implementation of national standards of quality nationwide.

These priorities triggered the need for additional innovative approaches in 2017 for the second phase
(Phase 2) of the QI program in Cambodia. At the time, Cambodia did not have a national standard of quality, a standard tool for laboratory assessment, or a law to enforce laboratory quality in the public or private sectors. To address this issue, the Cambodia Laboratory Quality Management System Checklist for Accreditation (CamLQMS) was developed and adopted by the MOH Bureau of Medical Laboratory Services (BMLS) as a tool for national auditors to assess laboratory quality during on-site performance audits. The CamLQMS tool was modeled on the WHO-AFRO SLIPTA tool, which is aligned with the ISO 15189 standard. Between July 2017 and April 2019, the Phase 2 QI program used the CamLQMS tool to track laboratory progress toward meeting ISO 15189 accreditation standards, while using a combination of training, on-site and remote mentorship, and advocacy.

Phase 2 of the QI program directed a set of technical approaches and interventions at both the national and regional levels to strengthen the interconnectivity and collaboration between laboratories in the tiered laboratory system for improved public health and clinical functions. At the national level, the program worked to address gaps in the legal and regulatory framework and documentation concerning the establishment of national quality and safety standards. At the facility level, primary activities encompassed the design and delivery of job-specific, competency-based education and training to quality assurance officers (QAOs) and laboratory managers selected from the Phase 1 cohort of the 12 national and regional clinical laboratories. Facility-based staff were trained in operational quality management and provided regular mentoring support through on-site technical assistance and telementoring consultations. In January of 2018, I-TECH Cambodia partnered with MOH-BMLS to conduct a baseline CamLQMS assessment of participating laboratories, followed by a national dissemination meeting to discuss findings and develop recommendations. These recommendations included a series of 11 training workshops to improve the LQMS operational practices of QAOs and laboratory managers and to eliminate deficiencies identified during the audits. LQMS trainings were designed using adult learning principles and accepted pedagogy to improve learner comprehension and competency through a combination of theoretical and practical learning methods oriented toward health professionals. These trainings consisted of large-group formal instruction interspersed with several focused and interactive sessions over 2–5 days, as well as smaller laboratory-based training workshops held regionally to emphasize practical proficiency in technical skills. Although subject matter content focused primarily on quality assurance and operational management, laboratory managers and QAOs were also provided with organizational leadership skill-building activities.

Phase 2 of the QI program also included close mentorship of laboratory staff by trained laboratory quality mentors. As described previously, mentors were technically experienced laboratory professionals, trained in QI, proficient in both English and Khmer, and employed in the project full time. Mentors periodically visited laboratories to deliver individualized training and coaching to each of the 12 laboratories and to address the gaps identified in the baseline audits; however, for Phase 2 of the project, the majority of mentorship and supportive coaching reached laboratories remotely using technologies such as Zoom, WhatsApp, and Facebook Messenger.

Modeled on the ECHO project but adapted independently by the project team for laboratory mentoring, Zoom video conferencing technology was used to connect with the cohort of laboratories weekly (though often 2 or 3 times per week, on-demand) in a community of practice environment. Weekly training sessions followed a structured training schedule designed over a period of 16 months. This schedule was organized into weekly topics and followed a format of teacher presentation, laboratory presentation, question and answer sessions, and action items for the following week. Time for peer networking was also provided, and conversations on Zoom often carried over into other platforms such as Messenger and WhatsApp in the days following each session. Remote training and mentoring sessions were designed to reach more geographically dispersed laboratory professionals without the limitations of resource-intensive travel, thus improving the cost-effectiveness of activities. Through the use of Zoom Pro accounts, project staff were able to schedule meetings for up to 100 participants for up to 2 hours, providing visual presentations and video demonstrations, with the added benefit that each session was recorded and available for later review by participants. These trainings were designed primarily for QAOs and laboratory managers; however, all laboratory staff were welcome and many additional staff also attended the weekly sessions, with each session recording up to 28 participants from the combined group of laboratories. Importantly, the program enjoyed strong engagement from MOH-BMLS, which was involved in all project planning, implementation, and monitoring including all formal training sessions, workshops,
and audits. This involvement was essential, ensuring the continuity and sustainability necessary for the program to be replicated within other laboratories once current funding had ended. Following this 16-month period of training and mentoring, a second CamLQMS audit was conducted and these data, along with an assessment of activity outputs, are the foundation of this study and program evaluation.

**METHODS**

*Research Methodology*

This study evaluated the outputs and outcomes of Phase 2 activities in the 12 participating laboratories during the evaluation period of January 2018 to April 2019 between program audits. Our evaluation used an uncontrolled longitudinal study to assess changes in LQMS compliance to international standards between baseline and endpoint measurements. A cross-sectional analysis was then used to compare postimplementation LQMS performance and conformity of intervention laboratories to a select group of nonintervention laboratories. Data management and basic descriptive statistics for all evaluation methods were performed using Microsoft Excel for Office 365. All complex calculations of statistics and hypothesis testing were performed using STATA 14 statistical software.

**Quantifying Mentoring and Training Activity Outputs**

For the description and enumeration of activity outputs, this study used monitoring and evaluation records collected by the project team. Outputs of interest, as listed in the logic model in Table 1, included (1) the number of trainings attended by personnel of the intended audience per laboratory, (2) the number of days of on-site mentorship provided to each laboratory, and (3) the amount of video conference training and mentoring time attributed to individual laboratory personnel during the evaluation period. Data sources included (1) attendance records for the 11 completed trainings, (2) project team member reports, (3) Zoom meeting records extracted from project team member’s Zoom accounts (reviewed to match laboratory and position details of meeting participants to their Zoom user names), and (4) supplementary records of remote mentoring sessions conducted via Messenger and WhatsApp from mentors. Datasets from each of these data sources were organized into separate spreadsheets for review and descriptive analysis. Attendance records for all 11 training events were organized by meeting date, and participant data were analyzed for each

| TABLE 1. Calculations of the 3 Primary Activity Outputs and the Cambodia Laboratory Quality Management System Audit Score Achieved Within the Evaluation Period |
|-------------------------------|-------------------|-------------------|-----------------|-----------------|
| No. Completed Trainings of Intended Participants (Total No. Participants) | Mentor Time on Site per Laboratory, Days | Video Conference Participation Time, Minutes | Audit Score Difference, % |
| Lab A | 19 (25) | 9 | 3,766 | 9 |
| Lab B | 21 (26) | 10 | 5,855 | 13 |
| Lab C | 24 (25) | 10 | 2,742 | 15 |
| Lab D | 23 (25) | 10 | 6,320 | 32 |
| Lab E | 25 (29) | 13 | 9,302 | 37 |
| Lab F | 24 (37) | 13 | 9,664 | 31 |
| Lab G | 23 (24) | 9 | 6,800 | 26 |
| Lab H | 21 (27) | 13 | 5,290 | 28 |
| Lab I | 28 (36) | 13 | 7,210 | 17 |
| Lab J | 22 (24) | 8 | 4,434 | 28 |
| Lab K | 22 (26) | 10 | 8,675 | 15 |
| Lab L | 22 (26) | 12 | 2,263 | 7 |
| Group mean ± SD | 23 ± 2 (28 ± 4) | 11 ± 2 | 6,027 ± 2,454 | 21 ± 10 |
Quantifying External Audit Outcomes

We used the CamLQMS checklist for accreditation as the primary outcome measure to determine the performance of each laboratory’s quality management system before and after the training and mentoring program interventions. The CamLQMS checklist was divided into 12 sections of laboratory quality with a total of 117 questions regarding whether a particular standard was met, and each question was assigned a numerical value that contributed to the audit score within each section and in the whole (Table 2). Mentored laboratories completed baseline CamLQMS audits in January of 2018 and outcome assessments in April of 2019. Additionally, a control group of representative public laboratories that did not receive LQMS training or mentoring (nonmentored/nonintervention) was selected for a cross-sectional comparison. Control laboratories were selected by MOH-BMLS as the nearest in comparable capacity in terms of the complementary package of activities and services, although these facilities differed significantly from mentored laboratories in terms of baseline level of training and number of staff. Laboratory audits of mentored and nonmentored facilities were conducted by 3 teams of auditors who were trained by the project team to assess facility conformity and nonconformity to the CamLQMS checklist. Each 4-person auditing team was led by a lead auditor and included at least 1 MOH-BMLS representative. During the audit process, the lead auditor asked each of the 117 questions of the laboratory in series, and the team reviewed responses at the end of each audit to determine whether the requirements of each question were met partially or in full, indicated by “yes,” “no,” or “partial.” Questions that were not applicable to a laboratory due to individual requirements or organizational complexities of the facility served were answered with “NA.” After completion of all audits, the 3 team leads reviewed all
audit data together to identify any recording error, bias or inconsistencies in scoring methodologies between teams, and moderated audit point allocations accordingly. Audit scores were calculated as a percentage of the total value of checklist items for each section and overall for each laboratory.

A Wilcoxon signed-rank test for nonparametric comparison for paired samples was performed to determine the strength of the difference between 2018 and 2019 audit scores of mentored laboratories for each section and summary overall. Nonparametric statistics were selected to maintain consistent assumptions of normality between the groups of small sample size. Mean audit scores and standard deviations were calculated in each of the 12 sections for visual comparisons between laboratory groups, and all sections with statistically significant differences in scores between years were documented with the level of significance. The percent change in overall audit scores in each section was calculated to present the magnitude of change visually, and these percent differences were used as the primary variables for a Spearman’s rank correlation assessment of the relationship strength between audit score improvement and laboratory personnel participation time in Zoom activities. An assessment of the statistical difference between audit scores of mentored LQMS laboratories and nonmentored, non-LQMS comparison laboratories was performed using the Wilcoxon rank-sum test for 2 independent samples. Comparisons were made for overall audit scores and scores for individual audit sections, and all sections with statistically significant differences between groups were again documented.

### RESULTS

Table 1 shows the outputs for each measured program activity and the corresponding increase in camLQMS audit score as the direct program outcome and reveals an output of 274 (mean=23 ± 10) target personnel trainings, 72,321 (mean=6,027 ± 2,454) minutes of video conference training, and 130 (mean=11 ± 2) visits to laboratories, resulting in an average positive percent difference of 21 ± 10% between the 2018 and 2019 overall audit scores. Video conference participation time was calculated from a sample size of 153 Zoom meetings with traceable usage reports out of a total of 261 meetings identified from supplemental mentor reports and program activity calendar entries. In terms of staff inputs, formal training and video conference activities included 2 primary mentors, 2 mentor trainees, the country project coordinator, and 3 laboratory systems technical and
senior technical specialists. Additionally, several MOH officials from BMLS and the National Institute of Public Health participated in formal trainings and in numerous video conference activities.

A Wilcoxon signed-rank test indicated that overall audit scores for mentored laboratories in 2019 were significantly higher (median score=57%) than overall audit scores for the same laboratories in 2018 (median score=40%, z=3.06, P=.002). In a comparison of scores for individual audit sections between years, the Wilcoxon signed-rank test indicated that mean 2019 scores for 11 out of 12 audit sections improved significantly (P<.01), with “information management” being the exception, which had been maintained but not significantly improved from an already high performance level at baseline (Figure 1). A cross-sectional comparison of the 2019 audit performance of mentored laboratories with the sample of nonmentored laboratories showed a large contrast in scores between groups (Figure 1) and by section (Figure 2).

A 2-sample Wilcoxon ranked-sum (Mann-Whitney) test indicated that overall audit scores for mentored laboratories in 2019 (median=57%) were significantly higher than overall audit scores for nonmentored laboratories (median=23%) in the same year (z=3.96, P=.0001). Mann-Whitney tests comparing individual audit sections similarly revealed significant differences in 11 of the 12 sections (P<.001) between intervention and nonintervention laboratories, with “information management” again being the exception, which was significantly different at the P<.01 level. In terms of percent difference in mean section scores between groups, “client management and customer service” as well as “occurrence management and process improvement” demonstrated the largest differences of 57% and 52% between groups. “Information management” again showed the smallest percent difference (27%) between groups. In an assessment of the relationship between mean audit score differences from 2018 to 2019 and the amount of participation time by individual laboratories in Zoom video conference training, a Spearman’s rank correlation showed a strong monotonic relationship between the 2 variables (r_s=0.59, P=.04) with significant certainty.

DISCUSSION
The quality audit scores of laboratories participating in this program improved significantly following implementation of the training and mentoring activities, demonstrating that the QI program achieved its intended effect. Laboratory performance from mentored sites was significantly higher in all measured categories of quality management than in laboratories with no training or mentoring support, and this study clearly showed a positive correlation between laboratory QI and participant contact time with trainers and mentors via remote mentoring. The strong correlation between remote mentoring through video conference calling and improved audit scores indicates

FIGURE 1. Overall 2018–2019 Cambodia Laboratory Quality Management System Checklist for Accreditation Audit Scores for Mentored Public Hospital Laboratories

- The dashed line represents the average audit score for nonmentored laboratories (2018 audit data not available).
Remote mentoring is an effective QI tool and also presents a cost-effective alternative to on-site mentoring, which requires frequent travel to remote, hard-to-reach laboratories. A 1-week site visit from Phnom Penh to Kratie, for example, costs approximately US$398 for travel, lodging, and per diem for local mentors, but a Zoom Pro account can cost as low as US$45 annually. Although further studies are needed to evaluate the cost-effectiveness of remote versus on-site telementoring and other inputs such as on-site training or mentorship, our results suggest a notable cost-benefit of telementoring for LQMS improvement compared with on-site training. Remote mentoring has the further benefit of providing on-demand professional support and networking. A qualitative study of the remote mentoring program in Cambodia identified a number of recurring themes of benefits identified by participants, including that additional remote training reinforced concepts and provided peer learning opportunities and on-demand guidance; however, laboratories strongly preferred a more structured training format in the local language if online training was used. The use of video conferencing technologies for medical education and consultation shows promise as a tool to create communities of practice between laboratorians and other health practitioners in the future, a practice that will prove all the more valuable during the COVID-19 pandemic, given that online platforms have become the primary means of accessing professional training and consultation for many medical professionals.

Notably, although attendance in formal, in-person training and the number of on-site visits were semi-controlled for variation and therefore could not be tested for a relationship to LQMS improvement, the relationship is expected. In particular, the content of the program’s formal training curriculum is reflected in several individual audit sections that demonstrated major improvements. Topics such as “documents and records,” “management review,” “occurrence management,” and “process control” received particular focus in formal trainings, and thus coincided with superior program outcomes. Of note, this program chose to deprioritize the topics of “information management” and “facilities and safety” due to topical overlap with other ongoing national training programs. During site visits, mentors worked closely...
with laboratories on specific technical needs such as improved use of quality indicators (“occurrence management”), quality control testing (“process control”), “equipment” verification, and “corrective action.” In later stages of the implementation period, mentors coached personnel on internal auditing in preparation for the second round of CamLQMS audits. This close mentoring approach was predicted to have contributed to the program outcomes; however, further research is needed to isolate the impact of our program’s site visit and formal training models from that of remote mentoring.

Notable limitations and recommendations for future research are as follows. First, the CamLQMS checklist for accreditation is designed to assess gaps in conformity within individual laboratories to drive improvement in each specific section of LQMS. Because individual audit sections have different maximum possible scores, and because some questions are inapplicable to certain laboratories, overall audit comparisons between laboratories and sections should be interpreted with caution. Nonetheless, these comparisons serve as a useful estimate of program activity efficacy. Second, because our cross-sectional comparison of final CamLQMS results in program laboratories with nonprogram laboratories does not compare rates of change between groups, further prospective studies are needed to compare the rate of improvement directly through a pre-post design with a larger sample of facilities. The comparison group is also limited in its usefulness because of critical differences in staff size and training input at baseline. Control laboratories had 4–8 employees per facility compared with approximately 9–33 employees in the participant group and did not benefit from Phase 1 inputs, which resulted in better audit scores at baseline for mentored laboratories and may have provided a learning advantage over nonintervention laboratories. Finally, because monitoring of Zoom session reports was incorporated late into the program evaluation, 2% of participation time in video conferencing could not be associated with or disassociated from individual laboratories, leaving the potential of misclassification bias against certain laboratories prone to using unidentified devices. Built-in user report tools such as within Zoom serve as an easy-to-use mechanism for monitoring and evaluation of remote training and mentoring programs; however, some effort is needed to ensure data quality as it is collected.

Conventional in-training programs are resource-intensive; however, as we have described here, programs that use remote mentoring and training tools such as Zoom can circumvent the need for frequent activities that are high-cost elements such as on-site workshops and coaching. In addition, they have the added benefit of reaching a larger audience than would otherwise be possible due to cost. These findings contribute to the limited body of qualitative research on remote mentoring as a practice, which describes success in QI outcomes due to improved accountability, collaborative problem solving, and increased awareness of the importance of laboratory quality. This evaluation strongly suggests that tiered laboratory systems in resource-limited countries such as Cambodia would benefit from national expansion of LQMS training and mentorship programs of a similar design, at scale, utilizing a structured curriculum and particularly remote training and mentoring methodologies.

### CONCLUSION

This program used a combination of training, mentoring, and advocacy to achieve rapid and significant outcomes in quality management system development. Participating laboratories performed significantly better in audits of performance and conformity than nonintervention laboratories, suggesting that an expansion of this methodology in Cambodia may benefit currently nonmentored laboratories significantly toward meeting national standards of quality. Although our findings indicate that significant progress has been made in meeting international standards of quality in laboratory practice, laboratories in the public sector and laboratories in Cambodia should continue to implement stepwise QI programs with an emphasis on improved connectivity of laboratories to professional training and mentorship for effective QI.

### Acknowledgments

We acknowledge the efforts of laboratory staff from all 12 participating laboratories in this project and give special thanks to Dr. Michael Noble for adapting the University of British Columbia’s quality management course for in-person delivery in Cambodia. We acknowledge and thank Cat Koehn, Joanna Mandelsohn, and Malin Chao for their project management support. We thank our colleagues from the Cambodian Ministry of Health, Bureau of Medical Laboratory Services, National Institute of Public Health, the U.S. Centers for Disease Control and Prevention in Cambodia, and World Health Organization Cambodia for their support of this project. We thank collaborators from the Diagnostic Microbiology Development Program, Fondation Mérieux, and Westline Education Group in Phnom Penh. This work was funded by the U.S. Department of Defense, Defense Threat Reduction Agency.

### Funding

This work was funded by the U.S. Department of Defense, Defense Threat Reduction Agency and the views of the authors do not represent those of the funder.

### Competing interests

None declared.

### REFERENCES

Remote Mentoring Using Video Conferencing for Laboratory Strengthening


Peer Reviewed

Received: March 21, 2020; Accepted: November 3, 2020; First published online: December 11, 2020


© Donovan 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-20-00128
Using Community Health Workers and a Smartphone Application to Improve Diabetes Control in Rural Guatemala

Sean Duffy, a Derek Norton, b Mark Kelly, c Alejandro Chavez, d Rafael Tun, e Mariana Niño de Guzmán Ramírez, a Guanhua Chen, b Paul Wise, d Jim Svenson f

Key Findings
- A smartphone application providing algorithmic clinical decision support enabled community health workers to improve diabetes control for a group of patients in rural Guatemala.

Key Implications
- Program managers should consider equipping community health workers with clinical decision support applications to enable task sharing for chronic disease management.
- Researchers should examine the efficacy of this approach for chronic diseases other than diabetes and compared to traditional models of care.

Resumen en español al final del artículo.

ABSTRACT

Background: The global prevalence of diabetes has nearly doubled since 1980. Seventy-five percent of patients with diabetes live in low- and middle-income countries, such as Guatemala, where health care systems are often poorly equipped for chronic disease management. Community health workers (CHWs) and mobile health technology have increasingly been applied to the diabetes epidemic in these settings, although mostly in supportive rather than primary roles in diabetes management. We sought to improve diabetes care in rural Guatemala through the development of a CHW-led diabetes program and a smartphone application to provide CHWs with clinical decision support.

Methods: We worked with our local partners to develop a program model and the smartphone application (using the CommCare platform) and to train CHWs. We recruited patients with type 2 diabetes living in rural communities. Program evaluation used a single-group, pre-post design. Primary outcomes were hemoglobin A1c and the percentage of patients meeting A1c goals compared with baseline. We also followed a variety of process metrics, including application reliability.

Results: Eighty-nine patients enrolled during the study period. The hemoglobin A1c percentage decreased significantly at 3 months (-1.0; 95% CI=-1.7, -0.6), 6 months (-1.5; 95% CI=-2.2, -0.8), 9 months (-1.3; 95% CI=-2.0, -0.6), and 12 months (-1.0; 95% CI=-1.7, -0.4). The percentage of patients with A1c <8% increased significantly at 3 months (23.6% to 44.4%, P=0.007), 6 months (22.0% to 44.0%, P=0.015), and 9 months (23.9% to 45.7%, P=0.03). CHWs and supervising physicians agreed with application medication recommendations >90% of the time.

Conclusion: Our results suggest that CHWs can safely and effectively manage diabetes with the assistance of a smartphone application and remote physician supervision. This model should be evaluated versus other standards of care and could be adapted to other low-resource settings and chronic diseases.

INTRODUCTION

The global prevalence of diabetes has increased dramatically over the past several decades, nearly doubling since 1980, from 4.7% to 8.5% of adults. In 2015, an estimated 5 million deaths and US$673 billion in health expenditures were attributable to diabetes, accounting for...
12.8% of global all-cause mortality and 12% of global health expenditures. Low- and middle-income countries (LMICs), where 75% of people with diabetes now live and 80% of deaths due to diabetes occur, have been especially hard hit by this global epidemic. In addition to limited or episodic care, resources are scarce; in one study, only 29.6% of patients in low-income countries were currently taking diabetes medications compared with 74% in high-income countries.

The World Health Organization (WHO) has advocated for the use of nonphysician health workers in the care of diabetes and other chronic diseases as a means to strengthen primary health care systems in LMICs. This approach is often referred to as task shifting, although there is a growing consensus that task sharing is a more appropriate framework given the difficulty of completely shifting highly complex clinical tasks to less extensively trained health care workers. Evidence is increasingly showing that sharing responsibilities with nonphysicians can improve access to care and patient outcomes for noncommunicable diseases. Community health workers (CHWs, also called health promoters)—lay people who are trained to carry out a variety of tasks and are often from or have a close connection to the communities they serve—are a common type of nonphysician health worker and are being increasingly utilized in health systems around the world, particularly in LMICs. Programs using CHWs for targeting diabetes care have shown improvements in glycemic control and other diabetes outcomes compared with standard care. In most of these programs, CHWs have played supportive roles (e.g., providing patient education or care coordination) rather than direct care roles.

With adequate training and support, CHWs have the potential to improve access to care and health care delivery. We evaluated the efficacy of trained CHWs enabled by a clinical decision support tool in directly providing elements of diabetes care with remote physician supervision, a novel approach for which little evidence exists at this time.

## METHODS

### Setting

This project occurred in the municipality of San Lucas Tolimán in the Western Highlands region of Guatemala, a middle-income country in Central America. A large majority of the population belongs to the Kaqchikel Mayan indigenous group. Poverty rates in this mostly rural municipality are high, with 91.1% of people living in poverty (<US$3 per day) and 29.4% in extreme poverty (<US$1.60 per day). A recent cross-sectional study conducted in San Lucas and other neighboring municipalities found a high prevalence of type 2 diabetes (13.8%) and prediabetes (13.8%).

Our local partner was the San Lucas Mission (SLM), a nonprofit organization associated with the Catholic parish in San Lucas Tolimán providing health care and other social services to the estimated 34,713 people living in the municipality. A community needs assessment conducted in the summer of 2016 as part of the planning process for this project found that diabetes care in the rural villages was generally fragmented or inaccessible, medications and supplies were often in short supply, and patients had very limited diabetes-related knowledge, particularly with regard to self-care. Government-run rural health outposts ostensibly provide basic care for diabetes and other chronic diseases, but in reality, they are inadequately staffed and supplied. In San Lucas, they were not a reliable source of care for patients. While patients could seek care in the private sector, it was often cost prohibitive, particularly for medications. These challenges contributed to ineffective treatment: Only 58% of patients reported taking diabetes medications regularly, and only 13% were meeting blood glucose (BG) targets. These findings reflect prior analyses reporting poor access to effective diabetes care in Guatemala, particularly for rural, indigenous populations.

SLM partners with local CHWs, known as promotores de salud (health promoters). These CHWs are recruited from the communities they serve, are bilingual in Spanish and Kaqchikel (the predominant Mayan language in this area), and generally have the equivalent of a US sixth grade education, affording basic literacy. General training for the CHW program occurs one weekend per month for 3 years and focuses on health prevention and early identification of patients for referral to a physician. The small group of leaders for the CHW program, called coordinators, are salaried. The other CHWs work mainly on a volunteer basis, but they receive a stipend per half day of work on dedicated health programs.

SLM had previously established an innovative and successful CHW-led childhood nutrition program enabled by mobile health technology. We sought to build on this foundation to create a sustainable rural diabetes program.

### Program Development

Program development was an iterative process that involved our local partners at all stages. We
first developed an overall model for the program, as outlined in Figure 1. In this model, health promoters meet with patients on a monthly basis. The promoters use a clinical decision support (CDS) application to guide each visit. Using data entered by the promoters, including point-of-care glycemic testing, the application provides recommendations on the titration of oral hypoglycemics, management of diabetes complications, self-care counseling, and referral to the supervising physician. After each visit, patient data are uploaded to a secure server and reviewed by one of the supervising physicians, who then communicates any changes in the treatment plan or additional recommendations to the promoters. In order to remove cost as a barrier to care, the diabetes program provides services and medications free of charge.

We recognized that the services provided by this program, while intended to be an improvement on the status quo, were by no means comprehensive. Guidelines for limited resource settings also deem insulin, antihypertensives, and other therapies as essential elements of diabetes care. However, resources were not available to implement a comprehensive chronic disease system. Rather, glycemic control through oral medications and lifestyle counseling was deemed the immediate focus, with additional components to follow with enhanced resources and a successful mobile platform proof-of-concept.

Development of Clinical Protocols and Procedures

We developed protocols for assessing glycemic control, titration of oral hypoglycemics, identification and management of diabetes complications, and patient counseling. We based these protocols on guidelines published by the American Diabetes Association (ADA), the International Diabetes Federation (IDF), and Guatemalan organizations. SLM medical director Dr. Rafael Tun was integral to this process and provided final approval for all protocols.

Assessment of Glycemic Control

We used point-of-care hemoglobin A1c (A1c) results as our primary measure of glycemic control based on recommendations from ADA and IDF. Studies have demonstrated the potential of this technology to improve diabetes care in LMICs. We utilized A1CNow point-of-care capillary blood analyzers. The A1CNow test produces results in 5 minutes and can be performed with minimal training, allowing for assessment of glycemic control by the CHWs during diabetes visits. Guidelines recommend checking A1c every 2–6 months depending on diabetes control and changes in medication. We checked A1c every 3 months for all patients during the study period to allow for more uniform evaluation of program efficacy.

We also employed monthly BG testing to titrate medications between A1c measurements,
In addition to providing algorithmic decision support to the CHWs, the application also served as a data collection tool and medical record.

**Medication Titration**

We selected metformin and glyburide (glibenclamide) as the oral medications in our medication titration protocol because of their long track records in diabetes care, availability in Guatemala, and affordability. Metformin is the first-line medication for all patients, consistent with established guidelines, with glyburide added as a second agent when glycemic targets are not met. For patients with an initial A1c of ≥9%, the algorithm calls for dual therapy (metformin and glyburide), as recommended by ADA and American Association of Clinical Endocrinologists/American College of Endocrinology guidelines. The titration algorithm accounts for 4 factors in making medication recommendations: glycemic control, current medication dose(s), adherence, and side effects.

**Identification and Management of Diabetes Complications**

We developed protocols for common and important diabetes complications and comorbidities, including hyper- and hypoglycemia, hypertension, coronary artery disease, chronic kidney disease, diabetic foot ulcers, and diabetic eye disease (Table 1). These protocols include recommendations for referral to the supervising physician and, in some cases (e.g., hypoglycemia), initial treatment delivered by CHWs.

**Application Development and Description**

We integrated the diabetes protocols into a smartphone application to provide algorithmic decision support to the CHWs. The application also served as a data collection tool and medical record. We designed the application in Spanish for smartphones and tablets running the Android operating system (Google LLC), the most common mobile operating system in Guatemala and globally. We used devices with quad core processors and 1 GB of RAM. While most patient visits were conducted at least partly in Kaqchikel, we did not translate the application to Kaqchikel based on feedback from the bilingual CHWs because Kaqchikel is primarily a spoken language and most CHWs are literate only in Spanish.

Earlier versions of the application used Enketo (Enketo LLC) web forms for the user interface and Ona (Ona Systems) for data storage and

### TABLE 1. Referral Protocols for Diabetes Complications and Comorbidities for Smartphone Application for Diabetes Care Program, Guatemala

<table>
<thead>
<tr>
<th>Routine Referrals (Within 1–2 Weeks)</th>
<th>Urgent Referrals (Within 1–2 Days)</th>
<th>Emergency Referrals (Same Day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage I hypertension (BP 140–160/90–100 mm Hg)</td>
<td>Stage II hypertension (BP 160–200/100–120 mm Hg)</td>
<td>Severe hypertension (BP ≥ 200/120 mm Hg)</td>
</tr>
<tr>
<td>Noninfected diabetic ulcer</td>
<td>Possibly infected diabetic ulcer, no signs of systemic infection</td>
<td>Fasting blood glucose undetectable high</td>
</tr>
<tr>
<td>Need for renal function testing</td>
<td>Worsening vision</td>
<td>Postprandial/random blood glucose undetectable high with mental status changes</td>
</tr>
<tr>
<td>A1c ≥ 9% despite maximal doses of metformin and glyburide for ≥3 months</td>
<td>FBG detectable, but ≥400 mg/dL</td>
<td>Hypoglycemia associated with altered mental status</td>
</tr>
<tr>
<td>A1c ≥ 9% for 3 consecutive checks</td>
<td>A1c ≥ 14%</td>
<td>Persistent hypoglycemia despite treatment in the field</td>
</tr>
<tr>
<td>A1c above glycemic target, but &lt;9% for 4 consecutive checks</td>
<td>Patient cannot tolerate minimum doses of metformin and/or glyburide</td>
<td>Current chest pain, high risk of CAD</td>
</tr>
<tr>
<td>Recent chest pain, moderate risk of CAD</td>
<td>Current chest pain, moderate risk of CAD</td>
<td>Possibly infected diabetic ulcer with signs of systemic infection</td>
</tr>
<tr>
<td>Blood in stool or possible melena</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient has other symptoms not addressed by the program protocols</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: A1c, hemoglobin A1c; BP, blood pressure; CAD, coronary artery disease; FBG, fasting blood glucose.
management, which was then transitioned to the CommCare platform (Dimagi, Inc.), the most widely used mobile platform among frontline health workers in LMICs. While both platforms allow for offline data collection and have branching logic capabilities, permitting the delivery of algorithmic clinical decision support, we transitioned to CommCare because it has more advanced capabilities for storing and modifying longitudinal data, includes robust database functions, and allows for application updates to be pushed to end-user devices. To maintain data security, we encrypted and password protected all the smartphones running the application. CommCare is also password protected and uses AES 256-Bit Symmetric Encryption, a HIPAA-compliant encryption standard.

Prior to deployment in the field, we tested application language, workflow, and user interface with the CHWs and reviewed the embedded clinical algorithms to ensure that the application provided appropriate recommendations. We continued to elicit feedback from the CHWs and update the application throughout the study.

CHW Training

CHWs were recruited for participation from the general rural health promoter program. We trained these CHWs in basic diabetes care (including medication management, diabetes self-care and lifestyle counseling, and the recognition and management of complications), protection of human subjects, and use of testing equipment (e.g., glucometers) and the application. We adapted training materials regarding diabetes self-care developed by 2 other Guatemalan organizations that work with indigenous populations, Wuqu’ Kawoq and Hospitalito Atitlán. To learn how to conduct finger-stick testing using glucometers and the A1cNow+ device, measure blood pressure using automatic cuffs, and accurately measure height, weight, and waist circumference, CHWs first viewed a demonstration of these skills and then practiced in small groups. Application training consisted of one-on-one practice with a facilitator to simulate a patient visit.

Total length of training was approximately 15 hours spread over several sessions. Dr. Duffy conducted the training sessions for the first several groups of CHWs. The coordinators of the CHW program led subsequent sessions. After receiving this initial structured training, CHWs were paired with one of the coordinators for patient care to continue supervised practice until they were able to complete a visit with minimal direction, a process which generally took 15 patient visits (approximately 9 hours).

Program Evaluation

Study Design

We used a single group, pre- and posttest design. Inclusion criteria for the program were established type 2 diabetes and age greater than 18 years. Exclusion criteria were insulin therapy, pregnancy, renal insufficiency (defined as estimated glomerular filtration rate [GFR] <30 mL/min/1.73 m²), and physician discretion.
Clinical Outcome Measures

Primary clinical outcomes were A1c and the percentage of patients with A1c ≤8% and meeting individual A1c goals compared with baseline. When A1c was higher than the detectable range of the A1CNow analyzer (displayed as “>13.0%”), we imputed these values conservatively as 13.1%. Secondary outcomes included BG, blood pressure, weight, body mass index (BMI), and waist circumference. When BG was higher than the detectable range of the Contour glucometers (>600 mg/dL, displayed as “HI”), we also imputed these values conservatively as 601 mg/dL.

We tracked the prevalence of medication side effects, change in medication dose, complications of diabetes and related referrals, and adverse events, with a focus on hypoglycemia (defined as BG <70 mg/dL) and hypoglycemia symptoms.

Behavioral Outcome Measures

We administered validated Spanish versions of 2 standardized questionnaires—the Diabetes Knowledge Questionnaire (DKQ)\(^41\) and the Summary of Diabetes Self-Care Activities (SDSCA)\(^42,43\)—in June 2018 to 2 subgroups of patients: patients enrolled in the past 3 months and those who had been participating for 6 months or more. We repeated questionnaires for patients in the newly enrolled group in January 2019.

Application-Specific and Process Outcomes

For each visit, we tracked whether the CHWs and the supervising physician agreed with medication recommendations provided by the application. We also tracked instances in which the application provided erroneous recommendations (as determined by the physician reviewing visit data). We administered a Spanish translation of the System Usability Scale, the most widely used standardized usability questionnaire, to all CHWs who had used the application. This scale results in a usability score from 0 to 100. We used a grading schema proposed by Bangor et al.,\(^44\) which rates usability scores less than 50 as “not acceptable,” those between 50 and 70 as “marginally acceptable,” and scores above 70 as “acceptable.” This usability survey also solicited written feedback about the application. Finally, we maintained detailed records of program costs in order to estimate the average cost per patient.

Patient Recruitment

Based on cases known to the CHWs, we estimated the number of patients with diagnosed diabetes in the rural communities of interest to be approximately 150. The CHWs recruited these patients for the program and we set an enrollment target of 100 patients, which reflected the resources and CHW capacity available for the program.

Statistical Analysis

We used R (The R Foundation) for analysis of program outcomes. We analyzed differences in continuous variables (e.g., A1c) using generalized additive mixed effects models (GAMMs) with the nonlinear smoothing function on time since program enrollment. For all health outcomes, baseline covariates of age, sex, and years since the...
participant’s diabetes diagnosis were included as standard main effects; the penalized regression splines were used on the longitudinal covariate of time since enrollment. Models also included subject-specific random intercepts and time-since-enrollment slopes. For the glucose outcome, whether the participant had been fasting at the time of measurement was also included in the models as a longitudinal main effect.

For the outcomes of A1c and glucose, values that were at the limit of detection were treated as a typical value in these GAMMs. In order to test if significant change in health outcomes from baseline occurred at 3, 6, 9, and 12 months after enrollment, bootstrapped confidence intervals were employed. Due to the censored nature of some of the A1c and glucose values, a sensitivity analysis of these outcomes was conducted using a Cox proportional hazard mixed effects modeling structure. Model diagnostics revealed a concern for heteroscedasticity in the glucose model. Refitting the model on the natural-log of glucose alleviated the issue, thus all reported glucose modeling results are from a model fitted to the natural-log of glucose.

The same GAMM structure already described was used to analyze A1c control (≤8%) and A1c goal attainment separately, with the appropriate model setup changes for the outcome being binary instead of continuous. Additionally, for A1c control/goal attainment, a pre-post study design was mimicked within the data by selecting each participant’s baseline value and their value closest to the 3, 6, 9, and 12 month follow-up period (within ±45 days, otherwise the observation at follow-up was considered missing). These pairs were then used to perform a McNemar test on the change in A1c control/goal attainment at these 4 follow-up times.

For DKQ and SDCA scores and medication doses, we used the Shapiro-Wilk test to determine normality. We then used 2-tailed t tests to assess differences in normally distributed variables and the Wilcoxon test for nonnormally distributed variables. We used a significance threshold (α) of 0.05 for all analyses.

**Ethical Oversight and Funding**

The program was reviewed and approved by the University of Wisconsin and Stanford University institutional review boards, as well as the SLM Health Program. All patients provided written informed consent after a bilingual CHW explained the study and risks and benefits of participation in the patients’ preferred language (Spanish or Kaqchikel). A seed grant from the University of Wisconsin Global Health Institute provided funding.

### RESULTS

#### Enrollment and Retention

Eighty-nine patients enrolled during the study period (February 2017 to June 2019), and 67 remained in the program at the end of this period (retention rate of 75.3%). Of patients who completed at least one follow-up visit, median follow-up time was 12.1 months (range 1.1–28.2, IQR 9.8). One patient died while participating in the program, 2 withdrew, and 11 were lost to follow-up. Eight patients were excluded after enrollment, 4 because of renal failure, 1 because of recurrent hypoglycemia while taking metformin alone, 1 because of hyperglycemia requiring insulin therapy, and 2 because of terminal illness.

Patients completed 920 visits (enrollment and monthly), 80.8% occurring at the designated central location and 19.2% in patient homes. Patients who remained in active follow-up completed 93.8% of possible visits, with all patients (including those who were excluded or lost to follow-up) completing 80.7% of possible visits.

#### Cohort Profile

Table 2 summarizes the baseline characteristics of enrolled patients, including place of diagnosis (a proxy for prior source of care) and medication use. Of note, a large majority (82%) of enrolled patients were women. Baseline glycemic control was poor, with a mean A1c (standard deviation [SD]) of 10.0% (2.5) and only 20% of patients meeting A1c treatment goals.

#### Clinical Outcomes

GAMM regression results are displayed in the Supplement. Age at baseline was significantly associated with A1c (β = −0.046, P = 0.002), natural-log glucose (β = −0.008, P = 0.003), systolic blood pressure (β = 0.569, P < 0.001), A1c control (OR = 1.05, P = 0.005), and A1c goal attainment (OR = 1.08, P < 0.001) but not associated with diastolic blood pressure, weight, waist circumference, or BMI. Baseline years since diabetes diagnosis was significantly associated with A1c (β = 0.073, P = 0.021), natural-log glucose (β = 0.018, P < 0.001), A1c control (OR = 0.90, P = 0.005), and A1c goal attainment (OR = 0.89, P = 0.025), but no other health outcomes. Fasting status was only in the glucose model.
and was significantly associated with natural-log glucose ($\beta = -0.357, P < 0.001$). Sex was not associated with any of the health outcomes examined.

Time since program enrollment was significantly associated with the outcomes of A1c, natural-log glucose, weight, and BMI (all $P < 0.001$), with non-linear behavior between times since enrollment and these outcomes. Figures 2 to 4 show the estimated behavior over time for A1c, glucose, and weight; BMI and weight results were very similar to one another, as expected, and only the weight figure is shown. Both A1c and glucose were estimated to decrease up to around 6 months, and then slowly rise back towards baseline values afterwards. However, the sparsity of observations after 1 year resulted in increased uncertainty in the estimated trend after this point. Both weight and BMI were estimated to increase up to about 6 months after baseline, then to slowly decrease afterwards. After 1 year, the uncertainty in the estimation of the trend increased greatly.

Based on the bootstrapped results for changes from baseline at 3, 6, 9, and 12 months (Table 3), A1c displayed significant reductions from baseline at all 4 intervals, with the greatest estimated reduction at 6 months (1.45 A1c % mean reduction), but still a 1-point reduction estimated at 1 year after enrollment. Natural-log glucose displayed significant reductions from baseline at 3 and 6 months, with the greatest estimated reduction at 6 months (0.135 natural-log mL/dL mean reduction; 22.4 mL/dL reverse transformed for a typical subject in the data; see Table 3 footnote), but the reduction at 9 and 12 months was nonsignificant. Weight displayed a significant mean gain over baseline at 3 months (1.86 lb estimated gain), but no significant change from baseline at the other time points. BMI displayed a similar trend to weight.

GAMM results for A1c control/goal attainment (Figures 5 and 6) showed an initial trend towards increased attainment until approximately 6 months, followed by a trend back towards baseline. No significant association was found between time since enrollment and probability of A1c control/goal attainment. However, when mimicking a pre/post design and analysis for examining these outcomes at 3, 6, 9, and 12 months from baseline, time periods closer to baseline were associated with significant increases in the proportion with A1c control/goal attainment (Table 4), similar to the A1c continuous analyses above. For A1c control, significant proportion increases from baseline were detected at 3, 6, and 9 months after baseline ($P$-values < 0.034), but not at 12 months ($P$ = 0.121). For A1c goal attainment, significant proportion increases from baseline were detected at 3 and 6 months ($P$-values < 0.020), but not at 9 and 12 months ($P$-values > 0.114). However, for both outcomes and at all 4 follow-up periods, the raw proportion increased, ranging from a 17.1% to 22.0% increase in A1c control, and from 7.3% to 20.0% increase in A1c goal attainment.

Sensitivity analyses using Cox mixed effects models to handle the true censored nature of the A1c and glucose values had numerous issues with the assumptions of proportionality. The results of

A community health worker in Guatemala conducts a home visit with a diabetes patient using a smartphone application for clinical decision support. Credit: © 2017 Sean Duffy/University of Wisconsin
the A1c model confirmed the GAMM A1c results, with increased time since enrollment associated with decreased A1c values (a significant "increased hazard" of observing A1c at lower values). The glucose model did not display a significant association between glucose and time since enrollment. However, from the GAMM results, there appeared to be nonlinear behavior between glucose and time. The Cox model did not properly account for this nonlinearity, and the "fall then rise" nature of the trend paired with the assumption violations could obscure a true association. The GAMM results did not show any significant relationship between time since enrollment and blood pressure or waist circumference, and the bootstrapped results did not show any significant difference in these variables at the designated analysis time points. We also ran unadjusted analyses for all the models described above, which were consistent with the adjusted results in terms of statistically significant associations and the nonlinear forms between time since intervention and outcomes.

### Medication Titration, Side Effects, and Adverse Events

Median daily doses of metformin and glyburide increased significantly (all \( P \leq 0.02 \)) from pre-enrollment to first recommendation at enrollment visit (500 to 1,700 mg and 0 to 2.5 mg) and from enrollment visit to last visit (1,700 to 2,550 mg and 2.5 to 5 mg). Patients taking metformin reported typical gastrointestinal side effects during 6.7% of visits. Side effects were significant enough to warrant a dosage reduction per the titration protocol during 3.9% of metformin-exposed visits (29.9% of metformin-exposed patients, 0.5 events per patient-year of therapy). There were 11 episodes of documented hypoglycemia (BG <70 mg/dL). Glyburide dosage was reduced due to hypoglycemia symptoms or documented hypoglycemia for 7.8% of glyburide-exposed visits (36.1% of glyburide-exposed patients, 0.9 events per patient-year of therapy).

Nine of the 11 hypoglycemic episodes were mild and resolved with treatment by CHWs or at home. Two hypoglycemic episodes required hospitalization for management. Both episodes occurred in the same patient, who was taking metformin alone and also had concomitant severe acute illnesses at the time of the episodes.

### Complications of Diabetes

Forty-four patients (49.1%) were identified as having increased risk of chronic kidney disease. Of these patients, 35 (80%) underwent renal function testing. Mean (SD) GFR was 77.1 (34.7) mL/min/1.73 m^2. Twenty-six (74.3%) patients in this group had normal GFR (>60 mL/min/1.73 m^2), 5 (14.3%) had GFR 30–60 mL/min/1.73 m^2, and 4 (11.4%) had significantly reduced renal function with GFR <30 mL/min/1.73 m^2.

A total of 279 referrals were recommended by the application for one or more potential complications of diabetes, representing 30.3% of visits. Of these, patients accepted 134 (48.0%) referrals. Based on available referrals tracking data, we estimate that patients completed 50.0% of accepted referrals, representing 24.0% of all recommended

---

**TABLE 2. Baseline Characteristics of Patients Enrolled in a Rural Diabetes Care Program, Guatemala**

<table>
<thead>
<tr>
<th>Characteristic (N=89)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
</tr>
<tr>
<td>Mean age (SD), years</td>
<td>53.5 (13.3)</td>
</tr>
<tr>
<td>Sex, % female</td>
<td>82</td>
</tr>
<tr>
<td>Years since diabetes diagnosis, median (IQR)</td>
<td>4 (6)</td>
</tr>
<tr>
<td><strong>Place of diagnosis, %</strong></td>
<td></td>
</tr>
<tr>
<td>San Lucas Mission rural clinic</td>
<td>40</td>
</tr>
<tr>
<td>Private clinic</td>
<td>20</td>
</tr>
<tr>
<td>Nongovernmental organization hospital</td>
<td>16</td>
</tr>
<tr>
<td>Guatemalan Social Security clinic</td>
<td>12</td>
</tr>
<tr>
<td>Government clinic</td>
<td>6</td>
</tr>
<tr>
<td>Other</td>
<td>6</td>
</tr>
<tr>
<td><strong>Medication use, %</strong></td>
<td></td>
</tr>
<tr>
<td>Taking any diabetes medication</td>
<td>82</td>
</tr>
<tr>
<td>Metformin</td>
<td>71</td>
</tr>
<tr>
<td>Glyburide</td>
<td>30</td>
</tr>
<tr>
<td>Glimepiride</td>
<td>3</td>
</tr>
<tr>
<td>Natural remedies</td>
<td>18</td>
</tr>
<tr>
<td><strong>Clinical measures</strong></td>
<td></td>
</tr>
<tr>
<td>Mean hemoglobin A1c (SD), %</td>
<td>10.0 (2.5)</td>
</tr>
<tr>
<td>Proportion with A1c at goal, %</td>
<td>20</td>
</tr>
<tr>
<td>Mean body mass index (SD), kg/m2</td>
<td>26.7 (4.6)</td>
</tr>
<tr>
<td>Mean blood glucose (SD), mg/dL</td>
<td>237 (126)</td>
</tr>
</tbody>
</table>

Abbreviations: IQR, interquartile range; SD, standard deviation. *Does not include natural remedies."
referrals. Figure 7 lists referrals by indication. Renal function testing was by far the most common indication for referral (50.5% of all referrals). Our clinical algorithms call for repeat referrals for renal function testing until completed for patients for whom it is indicated, contributing to the high number of referrals for this indication.

One patient died while participating in the program. The probable cause of death was myocardial infarction. This patient had well-controlled diabetes on metformin alone and had not reported symptoms of myocardial ischemia or other complications prior to their death.

Behavioral Outcomes
Thirteen patients who had been in the program for 6 months or more and 11 patients who had enrolled in the past 3 months completed the DKQ
and SDSCA. DKQ scores did not vary between the 2 groups, with a mean score of 13 for both (P=1). Of the newly enrolled patients, we were able to repeat the DKQ 6 months later for 5 patients. There was no significant difference between baseline and follow-up scores (mean 13 vs. 12.8, P=1).

Patients who had been enrolled for at least 6 months had higher average SDSCA scores (scored 0 to 7, with 7 being optimal) in several self-care categories compared with newly enrolled patients (see Table 5). However, only differences in foot care and dedicated exercise were statistically significant, with dedicated exercise scores being better in the newly enrolled group. We obtained follow-up SDSCA scores for 5 patients in the recently enrolled group 6 months after the initial questionnaire, which did not show any statistically significant improvements.

### Application-Specific and Process Outcomes

CHWs and the reviewing physician agreed with medication recommendations given by the application for 90.9% of visits. During 53 visits (5.8%), medication recommendations were altered by the CHWs after remote consultation with a physician. The reviewing physician changed medication recommendations based on data review after 30 visits (3.3%). There were 4 cases in which the application made inappropriate recommendations or malfunctioned. In each of these cases, patient treatment was corrected through direct communication between the supervising physician and the CHWs and future errors were prevented through application updates.

Twenty-one CHWs completed the System Usability Scale survey in January 2019. The mean score for fully completed surveys was 61.3 (range 27.5–87.5) and the mean composite score (accounting for responses from partially completed surveys) was 62.1. Subgroup analysis of scores above and below the predefined “acceptable” threshold of 70 showed that CHWs who rated application usability above 70 (n=4) were younger (mean age 32.0 vs. 42.2 years), more educated (mean 10.2 vs. 5.8 years of education), used smartphones more often (median use daily vs. once weekly), and had greater experience with the diabetes application (median use 11–15 times vs. less than 5 times) on average than those with scores 70 or below (n=12). Fourteen CHWs provided written subjective feedback on how the application could be improved. Common recommendations for improvement were to make the application faster and more responsive, reduce the number of questions and simplify language, and increase the amount of practice that CHWs had with the application.

We estimated a program start-up cost of US $3,940 for 100 patients, with continuing costs of US$118 per patient, per year (Table 6).

### DISCUSSION

Our results from the development and implementation of this program suggest that CHWs enabled...
### TABLE 3. Bootstrapped Results for Change Since Baseline in Outcomes Among Patients Enrolled in Rural Diabetes Care Program, Guatemala

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Time Since Baseline</th>
<th>Estimated Change</th>
<th>95% CI Lower Bound</th>
<th>95% CI Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A1C, %</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 months</td>
<td>–1.04(^a)</td>
<td>–1.68</td>
<td>–0.559</td>
<td></td>
</tr>
<tr>
<td>6 months</td>
<td>–1.45(^a)</td>
<td>–2.19</td>
<td>–0.813</td>
<td></td>
</tr>
<tr>
<td>9 months</td>
<td>–1.32(^a)</td>
<td>–2.01</td>
<td>–0.636</td>
<td></td>
</tr>
<tr>
<td>12 months</td>
<td>–1.03(^a)</td>
<td>–1.73</td>
<td>–0.385</td>
<td></td>
</tr>
<tr>
<td><strong>Glucose, natural-log mL/dL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 months</td>
<td>–0.104(^a)</td>
<td>–0.199</td>
<td>–0.0244</td>
<td></td>
</tr>
<tr>
<td>6 months</td>
<td>–0.135(^a)</td>
<td>–0.232</td>
<td>–0.0366</td>
<td></td>
</tr>
<tr>
<td>9 months</td>
<td>–0.0909</td>
<td>–0.163</td>
<td>0.00368</td>
<td></td>
</tr>
<tr>
<td>12 months</td>
<td>–0.0677</td>
<td>–0.166</td>
<td>0.0175</td>
<td></td>
</tr>
<tr>
<td><strong>Glucose, b mL/dL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 months</td>
<td>–17.5(^a)</td>
<td>–36.3</td>
<td>–5.00</td>
<td></td>
</tr>
<tr>
<td>6 months</td>
<td>–22.4(^a)</td>
<td>–39.2</td>
<td>–5.20</td>
<td></td>
</tr>
<tr>
<td>9 months</td>
<td>–15.4</td>
<td>–27.3</td>
<td>1.70</td>
<td></td>
</tr>
<tr>
<td>12 months</td>
<td>–11.6</td>
<td>–30.7</td>
<td>3.28</td>
<td></td>
</tr>
<tr>
<td><strong>Systolic BP, mm Hg</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 months</td>
<td>0.375</td>
<td>–5.02</td>
<td>1.08</td>
<td></td>
</tr>
<tr>
<td>6 months</td>
<td>0.75</td>
<td>–3.62</td>
<td>2.48</td>
<td></td>
</tr>
<tr>
<td>9 months</td>
<td>1.13</td>
<td>–2.84</td>
<td>3.87</td>
<td></td>
</tr>
<tr>
<td>12 months</td>
<td>1.5</td>
<td>–2.82</td>
<td>3.83</td>
<td></td>
</tr>
<tr>
<td><strong>Diastolic BP, mm Hg</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 months</td>
<td>–0.0678</td>
<td>–0.812</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>6 months</td>
<td>–0.189</td>
<td>–1.22</td>
<td>1.77</td>
<td></td>
</tr>
<tr>
<td>9 months</td>
<td>–0.467</td>
<td>–1.83</td>
<td>1.74</td>
<td></td>
</tr>
<tr>
<td>12 months</td>
<td>–0.877</td>
<td>–2.59</td>
<td>1.02</td>
<td></td>
</tr>
<tr>
<td><strong>Weight, lb</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 months</td>
<td>1.86(^a)</td>
<td>0.355</td>
<td>3.29</td>
<td></td>
</tr>
<tr>
<td>6 months</td>
<td>2.84</td>
<td>–0.0432</td>
<td>5.17</td>
<td></td>
</tr>
<tr>
<td>9 months</td>
<td>2.44</td>
<td>–1.67</td>
<td>4.88</td>
<td></td>
</tr>
<tr>
<td>12 months</td>
<td>1.67</td>
<td>–3.89</td>
<td>4.77</td>
<td></td>
</tr>
<tr>
<td><strong>Waist-circumference, cm</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 months</td>
<td>0.269</td>
<td>–0.474</td>
<td>1.28</td>
<td></td>
</tr>
<tr>
<td>6 months</td>
<td>0.51</td>
<td>–0.677</td>
<td>1.86</td>
<td></td>
</tr>
<tr>
<td>9 months</td>
<td>0.718</td>
<td>–0.632</td>
<td>1.99</td>
<td></td>
</tr>
<tr>
<td>12 months</td>
<td>0.896</td>
<td>–0.253</td>
<td>2.19</td>
<td></td>
</tr>
<tr>
<td><strong>BMI, kg/m(^2)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 months</td>
<td>0.372(^a)</td>
<td>0.0856</td>
<td>0.681</td>
<td></td>
</tr>
<tr>
<td>6 months</td>
<td>0.616(^a)</td>
<td>0.0477</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>9 months</td>
<td>0.639</td>
<td>–0.229</td>
<td>1.14</td>
<td></td>
</tr>
<tr>
<td>12 months</td>
<td>0.538</td>
<td>–0.674</td>
<td>1.1</td>
<td></td>
</tr>
</tbody>
</table>

*Continued*
by CDS technology can safely and effectively manage diabetes in rural Guatemala with remote physician supervision. Longitudinal analysis demonstrated significant improvements in the primary outcome of A1c, including at the predefined time points of 3, 6, 9, and 12 months after program enrollment. Statistically significant improvements in A1c ranged from 1.0% to 1.4%. These A1c improvements also meet the commonly used threshold of 0.5% for a clinically significant change in A1c.45,46

The proportion of patients with A1c ≤ 8% and meeting individualized treatment goals increased at each of these time points as well, with statistically significant increases at 3, 6, and 9 months and 3 and 6 months, respectively. However, it should be noted that significant covariates of age and years since diabetes diagnosis were not accounted for in these results. The GAMM analyses, which included these covariates, showed a trend in A1c control/goal attainment similar to that in the continuous A1c analysis, but the control/goal attainment trend did not meet statistical significance. Given that the continuous GAMM models showed significant improvements in A1c over time, the failure to detect statistically significant improvements in the adjusted binary A1c outcomes could have been a function of inadequate power.

The improvements in glycemic control associated with this program are similar to those reported for other CHW-led diabetes interventions in LMICs.21,47,48 A key difference from prior published interventions using CHWs in diabetes care is that rather than providing ancillary services, such as patient education, in support of traditional medical care, CHWs in our program were directly providing care: they assessed glycemic control, directed medication therapy, and identified potential complications with the assistance of mobile CDS technology. This approach is relevant for similar LMIC settings around the world, where health systems are faced with a rising tide of diabetes and other chronic diseases in the context of dire shortages of physicians, nurses, and other highly trained health workers.7,49

### Table 3. Continued

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Time Since Baseline</th>
<th>Est. Difference in Probability of Control/Goal</th>
<th>95% CI Lower Bound</th>
<th>95% CI Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1C control, A1C ≤ 8%</td>
<td>3 months</td>
<td>0.127</td>
<td>-0.0238</td>
<td>0.276</td>
</tr>
<tr>
<td></td>
<td>6 months</td>
<td>0.203</td>
<td>-0.0472</td>
<td>0.454</td>
</tr>
<tr>
<td></td>
<td>9 months</td>
<td>0.205</td>
<td>-0.0702</td>
<td>0.391</td>
</tr>
<tr>
<td></td>
<td>12 months</td>
<td>0.166</td>
<td>-0.0927</td>
<td>0.341</td>
</tr>
<tr>
<td>A1C goal, A1C ≤ subject goal</td>
<td>3 months</td>
<td>0.0588</td>
<td>-0.0096</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>6 months</td>
<td>0.0999</td>
<td>-0.0188</td>
<td>0.394</td>
</tr>
<tr>
<td></td>
<td>9 months</td>
<td>0.106</td>
<td>-0.0276</td>
<td>0.263</td>
</tr>
<tr>
<td></td>
<td>12 months</td>
<td>0.0929</td>
<td>-0.0359</td>
<td>0.195</td>
</tr>
</tbody>
</table>

Abbreviations: A1C, hemoglobin A1C; BMI, body mass index; BP, blood pressure; CI, confidence interval.

a *P* < 0.05.
b As these numbers are no longer on the scale of the regression, these values are specific to the type of subject the predictions were performed on (i.e., the values the other covariates are set at for prediction affect these numbers, unlike on the regression scale), which was the most common subject sex (female) and fasting value (true), median baseline age (54 years), and median years since diabetes diagnosis at baseline (4 years) for the subjects in analyses.

**Decision Support**

In general, the application provided reliable recommendations, with CHWs and the reviewing physician agreeing with the application-recommended medication dosing greater than 90% of the time. There were only 4 instances
in which the application provided incorrect recommendations compared with the established protocols. System Usability Scale surveys of the CHWs suggested marginally acceptable usability (mean score of 62.1). Subgroup analysis suggested that CHWs who had at least some high school level education, who used smartphones regularly, and who had more experience with the application found the application easier to use. While we elicited feedback from CHWs at all points of application development and deployment, this feedback was dominated by the coordinators of the CHW program, who were generally better educated and had more experience in conducting diabetes visits. Thus, increasing “rank and file”

FIGURE 5. Estimated Probability of Hemoglobin A1c Control (≤8%) Over Time After Enrollment in Rural Diabetes Care Program, Guatemala

![Graph showing probability of A1c control over time](image)

FIGURE 6. Estimated Probability of Meeting A1c Goal Over Time After Enrollment in Rural Diabetes Care Program, Guatemala

![Graph showing probability of meeting A1c goal over time](image)
CHW involvement in application development is one potential strategy to improve usability. CHWs also noted the tendency of the application to lag, negatively impacting usability. Our use of low-end Android devices likely accounts for this because we have found the application to work much faster on higher-performance devices. Fortunately, continued progress in smartphone development has meant that budget devices manufactured today are equivalent to flagship devices 2–3 years ago.50

WHO and other global health policy leaders have recognized the potential of mobile CDS tools to mitigate a lack of highly trained health care workers and supportive infrastructure and to improve the quality of care through the use of algorithmic protocols.38,51 These organizations have called for more rigorous evaluation of such mHealth interventions.51,52 Our experience in rural Guatemala adds to the evidence base supporting the use of mobile CDS to assist CHWs with chronic disease management and could be

### TABLE 4. Change in Proportion of Patients Meeting A1c Targets Among Those Enrolled in Rural Diabetes Care Program, Guatemala

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Time Since Baseline</th>
<th>N</th>
<th>Pre Control, %</th>
<th>Post Control, %</th>
<th>Proportion Change, %</th>
<th>McNemar P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1c control</td>
<td>3 months</td>
<td>72</td>
<td>23.6</td>
<td>44.4</td>
<td>20.8</td>
<td>.007</td>
</tr>
<tr>
<td></td>
<td>6 months</td>
<td>50</td>
<td>22</td>
<td>44</td>
<td>22</td>
<td>.015</td>
</tr>
<tr>
<td></td>
<td>9 months</td>
<td>46</td>
<td>23.9</td>
<td>45.7</td>
<td>21.8</td>
<td>.034</td>
</tr>
<tr>
<td></td>
<td>12 months</td>
<td>41</td>
<td>26.8</td>
<td>43.9</td>
<td>17.1</td>
<td>.121</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Time Since Baseline</th>
<th>N</th>
<th>Pre at Goal, %</th>
<th>Post at Goal, %</th>
<th>Proportion Change, %</th>
<th>McNemar P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1c at goal</td>
<td>3 months</td>
<td>72</td>
<td>16.7</td>
<td>31.9</td>
<td>15.2</td>
<td>.015</td>
</tr>
<tr>
<td></td>
<td>6 months</td>
<td>50</td>
<td>14</td>
<td>34</td>
<td>20</td>
<td>.016</td>
</tr>
<tr>
<td></td>
<td>9 months</td>
<td>46</td>
<td>15.2</td>
<td>28.3</td>
<td>13.1</td>
<td>.114</td>
</tr>
<tr>
<td></td>
<td>12 months</td>
<td>41</td>
<td>17.1</td>
<td>24.4</td>
<td>7.3</td>
<td>.55</td>
</tr>
</tbody>
</table>

Abbreviation: A1c, hemoglobin A1c.
adapted for diabetes management in similar LMIC settings. This approach could also be applied to other chronic diseases amenable to algorithmic care, such as hypertension. We will freely share the application to allow others to build upon our work.

Integration with the greater health system is integral to the success of mobile health applications. While our program does not directly interface with the government health system in Guatemala at this time, such regional or national partnerships would be essential for effective scale-up. TulaSalud, a nongovernmental organization working in the northern highlands of Guatemala, provides a model for effective scale-up in collaboration with the Ministry of Health.

---

**TABLE 5.** Comparison of Summary of Diabetes Self-Care Activities (SDSCA) Between New and Established Patients

<table>
<thead>
<tr>
<th>Measure</th>
<th>Patients Enrolled ≥6 Months (n=13)</th>
<th>Patients Enrolled &lt;3 Months (n=11)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy diet in the past week</td>
<td>7.0 [1.0]</td>
<td>6.0 [3.0]</td>
<td>.294</td>
</tr>
<tr>
<td>Healthy diet in general</td>
<td>7.0 [1.0]</td>
<td>6.0 [2.5]</td>
<td>.310</td>
</tr>
<tr>
<td>Eating fruits and vegetables</td>
<td>4.2 (2.1)</td>
<td>3.2 (2.5)</td>
<td>.319</td>
</tr>
<tr>
<td>Avoidance of high-fat foods</td>
<td>7.0 [1.0]</td>
<td>6.0 [0.5]</td>
<td>.088</td>
</tr>
<tr>
<td>Even distribution of carbohydrates</td>
<td>7.0 [0.0]</td>
<td>7.0 [0.0]</td>
<td>.755</td>
</tr>
<tr>
<td>Specific diet score</td>
<td>5.3 [1.1]</td>
<td>4.5 [1.1]</td>
<td>.088</td>
</tr>
<tr>
<td>General diet score</td>
<td>7.0 [1.0]</td>
<td>6.0 [2.8]</td>
<td>.336</td>
</tr>
<tr>
<td>Physical activity</td>
<td>7.0 [0.0]</td>
<td>7.0 [2.0]</td>
<td>.414</td>
</tr>
<tr>
<td>Dedicated exercise</td>
<td>0.0 [0.0]</td>
<td>0.0 [1.0]</td>
<td>.020</td>
</tr>
<tr>
<td>Exercise subscore</td>
<td>3.5 [0.0]</td>
<td>3.5 [1.5]</td>
<td>.424</td>
</tr>
<tr>
<td>Foot care</td>
<td>7.0 [1.0]</td>
<td>4.0 [5.5]</td>
<td>.047</td>
</tr>
<tr>
<td>Medication adherence</td>
<td>7.0 [0.0]</td>
<td>7.0 [0.0]</td>
<td>.849</td>
</tr>
</tbody>
</table>

*Values with parentheses represent mean (SD) and those with brackets represent median (IQR).*
and other health care organizations. Using the CommCare platform, they have developed and deployed mobile applications to assist CHWs in maternal and child health initiatives, and enable care coordination with the Ministry of Health, across a service area of 3.4 million people.

Medication Titration, Attenuation of Diabetes Control, and Medication Side Effects

It is possible that simply establishing consistent medication therapy through free provision of medications and regular follow-up, regardless of dose titration, accounted for improvements in glycemic control. Other studies of diabetes management in LMICs have shown marked improvements in A1c resulting from reconstitution of medication therapy, particularly when baseline A1c is high. However, 2 factors support the importance of the titration algorithm in our program. First, most patients (82%) reported that they were taking medications at the time of enrollment. Thus, subsequent improvements in glycemic control suggest that medication optimization and not merely initiation played a role for most patients. Secondly, median doses of metformin and glyburide increased significantly during the follow-up period.

Our data suggest possible attenuation of program effects on glycemic control over time. Although the reduction in A1c remained significant at 12 months after enrollment, A1c reduction peaked at 6 months and trended back towards baseline after this point. So-called “secondary failure” of hypoglycemic medications—a reduction in efficacy over time, particularly for glyburide and in patients with prior long-term, high-dose treatment—could also contribute to long-term attenuation of improvements in glycemic control.

Based on our data, the safety of the intervention was comparable to routine diabetes care delivered in other contexts. Patients experienced metformin side effects requiring dosage reduction at 3.9% of metformin-exposed visits. This outcome is comparable to clinical trials of metformin, which generally report a 5% prevalence of metformin intolerance. Thirty-six percent of glyburide-exposed patients experienced probable hypoglycemia symptoms or documented hypoglycemia, with a mean of 0.9 events per patient-year of therapy. None of these episodes were severe. Published estimates of the frequency of hypoglycemia attributable to glyburide and other sulfonylureas vary widely based on event definitions.

A prospective study of 383 patients that used a definition of hypoglycemia similar to ours (patient report of hypoglycemia symptoms or documented glucose measurement in the hypoglycemic range) found a similar prevalence (39%) and incidence (1.92 events per person-year) of hypoglycemia in patients taking sulfonylureas.

Diabetes Self-Care Counseling

While patients enrolled for at least 6 months had higher SDSCA scores than newly enrolled patients in several self-care categories, these differences were only statistically significant for foot care and dedicated exercise (with exercise scores actually better in the newly enrolled group). Additionally, improvements in A1c resulting from reconstitution of medication therapy, particularly when baseline A1c is high. However, 2 factors support the importance of the titration algorithm in our program. First, most patients (82%) reported that they were taking medications at the time of enrollment. Thus, subsequent improvements in glycemic control suggest that medication optimization and not merely initiation played a role for most patients. Secondly, median doses of metformin and glyburide increased significantly during the follow-up period.

Our data suggest possible attenuation of program effects on glycemic control over time. Although the reduction in A1c remained significant at 12 months after enrollment, A1c reduction peaked at 6 months and trended back towards baseline after this point. So-called “secondary failure” of hypoglycemic medications—a reduction in efficacy over time, particularly for glyburide and in patients with prior long-term, high-dose treatment—could also contribute to long-term attenuation of improvements in glycemic control.

Based on our data, the safety of the intervention was comparable to routine diabetes care delivered in other contexts. Patients experienced metformin side effects requiring dosage reduction at 3.9% of metformin-exposed visits. This outcome is comparable to clinical trials of metformin, which generally report a 5% prevalence of metformin intolerance. Thirty-six percent of glyburide-exposed patients experienced probable hypoglycemia symptoms or documented hypoglycemia, with a mean of 0.9 events per patient-year of therapy. None of these episodes were severe. Published estimates of the frequency of hypoglycemia attributable to glyburide and other sulfonylureas vary widely based on event definitions.

A prospective study of 383 patients that used a definition of hypoglycemia similar to ours (patient report of hypoglycemia symptoms or documented glucose measurement in the hypoglycemic range) found a similar prevalence (39%) and incidence (1.92 events per person-year) of hypoglycemia in patients taking sulfonylureas.
when we repeated this questionnaire with these newly enrolled patients after 6 months, there were no statistically significant improvements. Sample size was very small, including only 5 patients for repeated SDSCA questionnaires, so it is difficult to reach any conclusions on the effectiveness of CHW-delivered lifestyle counseling. However, the lack of significant change suggests that counseling may need to be intensified and optimized. Two other interventions in which CHWs and diabetes educators provided self-care counseling to indigenous Guatemalans with diabetes have reported significant improvements in glycemic control. Of note, both of these interventions were relatively intensive, with weekly visits in one intervention and mean counseling time of 10 hours over a 9-month period in the other. In contrast, visits in our program occur monthly and typically include approximately 10 minutes of diabetes self-care counseling.

Program Costs
The estimated cost of this program is less than that reported for a nurse-led diabetes program in Guatemala: US$118 versus US$220 per patient, per year. However, this program provided more comprehensive services, including insulin and hypertension treatment. The cost of our program is also comparable to data from a recent systematic review of the cost of diabetes treatment in LMICs, which reported average annual treatment costs ranging from US$29.91 to US$237.38 per person. 

Limitations
The primary limitation of this study is the lack of a control group. A future study comparing CHW-led care with physician, midlevel provider, or nurse-led care is necessary to determine the efficacy of our approach versus standard practice. Another limitation of our analysis was the substitution of inferred values for A1c and glucose when measurements fell outside the range of the measurement devices. This injects a degree of uncertainty into the calculated changes in mean A1c and glucose throughout the study. However, sensitivity analysis showed that changes in A1c were robust to this limitation in measurement. In addition, improvements in the proportion of patients meeting A1c goals were not affected by this measurement uncertainty, and this outcome supports the efficacy of the program in improving glycemic control.

Another issue inherent in A1c measurement is the effect of anemia, hemoglobinopathies, and other metabolic abnormalities. While hemoglobinopathies are rare in indigenous populations of the Americas, anemia (primarily iron-deficiency anemia) affects more than 20% of women of childbearing age in Guatemala. We did not screen subjects for anemia in this study, so we are unable to assess the potential effect of anemia on our results. However, the primary outcomes in this study were longitudinal with each subject acting as their own control, mitigating the potential effect of skewed A1c results due to anemia in our analysis.

Our study population was mostly women (82%). The “men’s health gap”—reduced health care utilization and poorer health outcomes among men compared to women—is an important global phenomenon. Other diabetes interventions in rural Guatemala have also struggled to recruit and retain men. The low participation levels of men are likely multifactorial, but in our experience the predominantly agricultural nature of men’s work in these communities, entailing long hours and lengthy travel to the fields, is a key factor. Despite offering home visits on weekends, we were unable to overcome these barriers. Further research is needed on how to improve outreach to men in rural Guatemala and similar contexts.

Due to a low referral completion rate, relatively few referrals for certain complications of diabetes (such as chest pain and vision problems), and lack of advanced diagnostic testing capabilities at the referral hospital, it is difficult to assess the accuracy and efficacy of our protocols for detection, management, and referral of potential diabetes complications. Although we did not have renal function testing available for our entire patient population to validate our algorithm for identifying patients at higher risk of renal impairment, 25.7% of patients who completed renal function testing had at least some degree of renal function impairment (GFR <60 mL/min/1.73 m²) and 11.4% had significant renal impairment (GFR <30 mL/min/1.73 m²). This is similar to the prevalence of decreased GFR in type 2 diabetics (22%) estimated from a large global study completed in 2006. Thus, even though we have testing data available for renal function, it is difficult to assess the effectiveness of our algorithm in identifying high-risk patients.

Finally, we designed this program and the CDS application to fit our specific context of rural
Guatemala and the specific resources and capacity of our local partner, which may make our findings less generalizable to other settings. While we are hopeful that others will be able to learn from our experience and to use the application, significant modifications may be required for our model to be used elsewhere.

### CONCLUSIONS

A novel CHW-led diabetes program enabled by mobile CDS technology led to improvements in diabetes control for a rural Guatemalan population. A task-sharing model using nonphysician health care workers assisted by mHealth applications holds promise for improving the care of diabetes and other noncommunicable diseases in LMICs, which represent a crucial health challenge of the 21st century. Further work is needed to determine the efficacy of this approach compared with standard care, to enhance the application to allow for the delivery of more comprehensive diabetes management, and to better support lifestyle changes through enhanced counseling and interventions to improve the nutritional environment.

### Acknowledgments:

We would like to especially acknowledge our partners in Guatemala—including San Lucas Mission Medical Director Dr. Rafael Tun and the community health workers, particularly José Vicente Maccario, Césia Castro Chutú, Olga Marina Ajacon Cuc, and Domínga Pic Salazar—who have been integral to this project. Dr. David Rabago was invaluable as a faculty mentor for Dr. Duffy, the primary author. Dr. Mindy Smith helped to edit and frame a first draft of the manuscript.

### Funding:

The University of Wisconsin Global Health Institute provided seed grant funding for this project. Dr. Duffy was also supported by the National Research Service Award Primary Care Postdoctoral Research Traineeship awarded by the National Institutes of Health, Health Resources Services Administration (HRSA grant T32 HP10010) during a portion of the time spent working on this project.

### Competing interests:

None declared.

### REFERENCES


Task Sharing Using mHealth to Improve Diabetes Control in Rural Guatemala


En español

Usando Trabajadores Comunitarios de la Salud y una Aplicación de Smartphone para Mejorar el Control de la Diabetes en una Zona Rural de Guatemala

Hallazgos claves

- Una aplicación para smartphones que brinda apoyo algoritmo a las decisiones clínicas permitió a los trabajadores comunitarios de salud mejorar el control de la diabetes en un grupo de pacientes en una zona rural de Guatemala.

Implicaciones claves

- Los administradores de programas deben considerar equipar a los trabajadores comunitarios de salud con aplicaciones de apoyo a la toma de decisiones clínicas para permitir el compartir de tareas para el manejo de enfermedades crónicas.

- Los investigadores deben examinar la eficacia de esta estrategia para enfermedades crónicas distintas de la diabetes y compararlo con los modelos tradicionales de cuidado médico.

Resumen

Antecedentes: La prevalencia mundial de diabetes casi se ha duplicado desde 1980. Setenta y cinco por ciento de los pacientes con diabetes viven en países de ingresos bajos y medianos, como Guatemala, donde los sistemas de atención médica a menudo están mal equipados para el manejo de enfermedades crónicas. Los trabajadores comunitarios de salud del país (TCS) y la tecnología de salud móvil se han aplicado cada vez más a la epidemia de diabetes en estos entornos, aunque principalmente en funciones de apoyo más que en el manejo directo de la diabetes. Buscamos mejorar la atención de la diabetes en las zonas rurales de Guatemala mediante el desarrollo de un programa de diabetes dirigido por las TCS y una aplicación para smartphones para brindarles apoyo en la toma de decisiones clínicas.

Métodos: Trabajamos con nuestros socios locales para desarrollar un modelo de programa y la aplicación para smartphones utilizando la plataforma CommCare y para capacitar a los TCS. Reclutamos pacientes con diabetes tipo 2 que vivían en comunidades rurales. La evaluación del programa utilizó un diseño pre-post de un solo grupo. Los resultados primarios fueron la hemoglobina A1c y el porcentaje de pacientes que alcanzaron los objetivos de A1c en comparación con el valor inicial. También seguimos una variedad de métricas de procesos, incluyendo la confiabilidad de la aplicación.
Resultados: Ochenta y nueve pacientes fueron inscritos durante el período de estudio. El porcentaje de hemoglobina A1c disminuyó significativamente a los 3 meses (−1,0; IC del 95%: −1,7 a −0,6), 6 meses (−1,5; IC del 95%: −2,2 a −0,8), 9 meses (−1,3; IC del 95%: −2,0 a −0,6) y 12 meses (−1,0; IC del 95%: −1,7 a −0,4). El porcentaje de pacientes con A1c < 8% aumentó significativamente a los 3 meses (23,6% a 44,4%, P = 0,007), 6 meses (22,0% a 44,0%, P = 0,015) y 9 meses (23,9% a 45,7%, P = 0,03). Los TCS y los médicos supervisores estuvieron de acuerdo con las recomendaciones de la aplicación para el uso de medicamentos más que el 90% del tiempo.

Conclusión: Nuestros resultados sugieren que los TCS pueden manejar la diabetes de forma segura y eficaz con la ayuda de una aplicación para smartphones y la supervisión médica remota. Este modelo debe evaluarse frente a otros estándares de cuidado médico y podría adaptarse a otros entornos de escasos recursos otras enfermedades crónicas.

Peer Reviewed

Received: February 13, 2020; Accepted: September 1, 2020; First published online: December 4, 2020


© Duffy 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-20-00076
Behavioral Insights Into Micronutrient Powder Use for Childhood Anemia in Arequipa, Peru

Jessica D. Brewer,a Julianna Shinnick,b Karina Román,c,d Maria P. Santos,a Valerie A. Paz-Soldan,a,d,e Alison M. Buttenheimb,f

Key Findings

- Negative experiences with health care providers or inconvenience at the time that micronutrient powder (MNP) use is initiated may discourage future MNP use.
- Mental models about nutrition can shape intentions to use MNP, and having too many choices can confuse caregivers.
- A single negative experience with MNP can form strong memories and discourage caregivers from giving MNP.

Key Implications

- Training for health care providers should encourage positive interpersonal interaction with caregivers during initiation of MNP because these interactions can have a lasting impact on MNP use.
- Education for caregivers should include counseling about potentially challenging side effects so that caregivers are prepared to work through them.
- In future programming, public health practitioners should consider encouraging caretakers to utilize well-timed cues to administer MNP.

ABSTRACT

Childhood anemia remains a significant driver of morbidity in low- and middle-income countries, including Peru. To identify behavioral challenges to using micronutrient powder (MNP) that is given to supplement children’s diets and prevent anemia, we applied a behavioral design approach to interviews and focus groups with 129 caregivers in Arequipa, Peru. We examined 3 key points in the decision-making process: accessing MNP through the health system; forming intentions to use MNP; and MNP use at the time of child feeding. Using the NUDGE (Narrow, Understand, Discover, Generate, Evaluate) approach, we identified the following behavioral barriers and facilitators: (1) caregivers’ experiences with health care providers shaped their motivation to access MNP; (2) caregivers felt accessing MNP at clinics was inconvenient and created hassle factors; (3) caregivers’ mental models about anemia prevention shaped MNP intentions and use; (4) caregivers’ salient negative experiences could have caused them to stop giving MNP; (5) caregivers forgot to give MNP if they did not have cues to remind them but could be prompted with salient cues; and (6) caregivers were affected by emotional, cognitive, and attentional factors during feeding that were difficult to anticipate. Our results, based on a behavioral design approach, suggest opportunities to adapt current messaging, counseling, and education around MNP use. Adaptations include providing culturally relevant messages, leveraging caregivers’ emotional and cognitive states, and encouraging small but impactful changes to feeding routines to address barriers to MNP use.

INTRODUCTION

Anemia in children can impair cognitive and motor function and cause fatigue and poor school performance.1–2 It is a significant public health issue, particularly affecting low- and middle-income countries (LMICs).3–4 In Peru, prevalence of anemia among children aged 6 months to 3 years was 43.6% in 2017.5 To combat this high prevalence, in 2014, the Peruvian Ministry of Health began distributing free micronutrient powders (MNPs) (or “chispitas”) to children aged 6 months to 3 years at public health care facilities during well child checkups, where caregivers receive guidance from health
Despite Peru having national guidelines on MNP use, adherence is low and child anemia prevalence remains high.

Behavioral economics analysis may offer novel insights into low adherence to MNP use despite its availability and promotion.

Behavioral Economics and Intervention Design

Several studies have identified social, psychological, and environmental factors that inhibit or enable MNP use; we extend that research here with an applied behavioral design approach, informed by behavioral economics, to understand the behavioral processes at play in MNP use. Behavioral economics—a field that sits at the intersection of economics and psychology—seeks to understand how common mental biases, heuristic thinking, and social forces shape decision making and behavior. A rich theoretical and empirical collection of literature from behavioral economics and related disciplines describes and characterizes how decision making often deviates from what rational actor or expected utility models would predict. Consistent findings in this interdisciplinary literature are that humans heavily rely on “rules of thumb” and mental shortcuts to make decisions, are given imperfect information, have time-inconsistent preferences, and have attitudinal and cognition constraints. These decisions are often not in their best long-term interest but satisfy their immediate needs and desires.

Bringing a behavioral economics perspective to the analysis of uptake of public health programs can help identify specific barriers to and facilitators of target behaviors that are not captured by other approaches. We define behavioral barriers as those factors arising from cognitive or psychological processes that reduce the likelihood of a target behavior being carried out; behavioral facilitators similarly increase that likelihood. Behavioral barriers and facilitators often operate separately from conscious cognition or awareness; one implication of this is that people’s statements about their intentions, motivations, and decisions around a behavior—particularly a complex or habitual behavior—may paint an incomplete picture of that behavior’s context. In recent years, innovative methods have emerged that map contextual data about a behavior (including field observations; interviews and focus groups with participants, stakeholders, and experts; and existing quantitative evidence and prior literature) to specific behavioral economics principles to uncover novel insights about barriers and facilitators that can inform intervention design. These methods and approaches have been widely used in a variety of global health settings and programmatic domains and in previous work on food choices and human nutrition.

Because the provision of MNP is an active intervention that requires multiple steps, sustained action over time, and the translation of intentions into behavior, behavioral economics may offer novel insights into low adherence to MNP use despite its availability and promotion. Interventions informed by behavioral economics have been used successfully in prior studies to improve maternal and child nutrition, from simple changes to the layout of school cafeterias and providing verbal cues for healthier choices in the United States to incentives and reminders to buy healthy foods in Madagascar that are designed to address specific behavioral barriers (for example, incentives to address procrastination and present-orientation, and stickers that deliver salient reminders at the point of purchase and consumption).

In this article, we advance our understanding of behavioral barriers to and facilitators of consistent...
MNP use for anemia prevention in Peru. Using previously collected contextual inquiry data, we applied a behavioral design approach to uncover novel insights about caregiver choices and actions related to giving MNP. These insights can inform counseling techniques used in MNP programs.

**METHODS**

**Setting**
We conducted our study in Arequipa, the second largest city in Peru, which has particularly high rates of childhood anemia.5 Despite ongoing efforts to address anemia, 44.5% of children aged 6 months to 3 years in Arequipa were diagnosed with anemia in 2016.44 The study was conducted in 8 of 29 districts in Arequipa, which accounted for more than half of the cases of early childhood anemia in the province according to unpublished sources from the local branch of the Ministry of Health.

**Data**
In 2017, we conducted 24 interviews and 12 focus groups with caregivers of children aged 6 months to 3 years. Caregivers were defined as adults who self-reported spending at least 5 days a week providing care for the child, whether they were the child’s biological parent or otherwise. This inclusion criterion was established as the only one for recruiting caregivers as we believe the primary caregiver is typically in charge of child feeding and thus the administration, or lack thereof, of MNP and so that we could obtain a range of caregiver experiences related to gender, caregiver age, child age, child history of anemia, caregiver-child relationship, and other factors. Caregivers were selected for interviews via convenience sampling in and around local health establishments and selected for focus groups through door-to-door recruitment in the neighborhoods surrounding the health establishment. The interview and focus group guides (Supplement 1) were developed to probe for caregiver experiences in obtaining MNP from health care providers and applying MNP to children’s meals, as well as other beliefs about anemia treatment and prevention. Interviews and focus groups were conducted by authors JDB, KR, and MPS in Spanish, audio recorded, and transcribed in Spanish. In total, we conducted individual interviews with 24 caregivers and 12 focus groups with 4 to 13 caregivers each, resulting in a total of 129 caregiver participants. The study team analyzed a subset of data from this parent study for the present analysis. More details about the data from the parent study are available from Brewer et al.15

**Analytical Approach**
We used the NUDGE (Narrow, Understand, Discover, Generate, Evaluate) approach to analyze behavioral barriers to and facilitators of MNP use among caregivers. NUDGE was developed to support the systematic and rigorous application of behavioral economics insights to intervention design,45 and is one of several published design approaches informed by behavioral economics and design thinking.33,46 The use of the term “nudge” is intentional; the approach generates intervention designs that are consistent with Sunstein and Thaler’s definition of nudges31:

*any aspect of the choice architecture that alters people’s behavior in a predictable way without forbidding any options or significantly changing their economic incentives.*

NUDGE includes 5 stages:
1. Narrow the focus of the analysis to a specific, relevant behavioral target
2. Understand the context of the behavior through inquiry into the decision-making process and related actions
3. Discover insights about barriers to and facilitators of the target behavior through structured matching of elements from contextual understanding developed in Stage 2 to core principles (cognitive biases and heuristic thinking) from behavioral economics
4. Generate intervention strategies and designs to address identified barriers
5. Evaluate those designs through iterative prototyping and trialing

In this article, we report the results from the Narrow, Understand, and Discover stages (Figure).

Building on a previous analysis utilizing the social-ecological model to identify factors that inhibited and enabled MNP use from Brewer et al.,15 we narrowed our point of inquiry to a defined behavioral target: the regular use of MNP during child feeding. We developed a rich understanding of the context around MNP use through repeated reading of focus group and interview data. To discover relevant behavioral insights, we first identified the key decisions and actions underlying the target behavior. Next, using a set of prompts about the cues, meanings, and alternatives
related to decision and action steps, we brainstormed barriers to or facilitators of each step. Each barrier linked the contextual understanding developed in the previous stage with 1 or more specific behavioral economics constructs (e.g., availability heuristic or present bias) to discover an insight about barriers and facilitators for the target behavior. Three examples of how a barrier or facilitator is discovered from contextual inquiry and behavioral constructs are shown in the Table.

This process yielded 121 barriers and facilitators related to MNP use, which were de-duplicated to...

### TABLE. Summary of Discovery of Behavioral Barriers/Facilitators to Micronutrient Powder Use Among Caregivers in Peru

<table>
<thead>
<tr>
<th>Decision-making Step</th>
<th>Prompt (Cue, Action, or Meaning)</th>
<th>Contextual Factor</th>
<th>Behavioral Construct</th>
<th>Barrier or Facilitator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accessing MNP</td>
<td>Cue: The perspectives of people with authority on MNP hold greater weight</td>
<td>Caregivers are more likely to make decisions about their children’s health when an expert gives them the information.</td>
<td>Authority bias</td>
<td>Negative interactions with medical professionals can cause caregivers not to give MNP.</td>
</tr>
<tr>
<td>Using MNP at the moment of child feeding</td>
<td>Meaning: Is the action uncomfortable or painful such that it is avoided?</td>
<td>Children often react negatively to taste of MNP, making feeding difficult or unpleasant for caregiver.</td>
<td>Negativity bias</td>
<td>If children refuse food supplemented with MNP, caregivers may stop giving it.</td>
</tr>
<tr>
<td>Using MNP at the moment of child feeding</td>
<td>Action: Caregivers abruptly discontinue MNP</td>
<td>Caregivers abruptly decide to stop giving MNP when their child has diarrhea or another side effect.</td>
<td>Hot-to-cold empathy gap</td>
<td>Visceral reactions when a child is sick can lead to a rapid choice to discontinue MNP.</td>
</tr>
</tbody>
</table>

Abbreviation: MNP, micronutrient powder.
68, synthesized to 21, and validated in an iterative process to ensure rigor and completeness. The validation process included in-depth discussions with authors KR and MPS, who are public health researchers involved in original data collection but who had not previously participated in this analytic process. This validation process resulted in the exclusion of 4 barriers (and related behavioral constructs) that did not align with the researchers’ experiences in the field. The validated list of 17 barriers and facilitators was further reduced to a set of 6 coherent, validated barriers to and facilitators of the target behavior that can inform intervention design. Importantly, the results of this analytic process are at the level of the barrier to or facilitator of the focal behavior; while each barrier/facilitator is informed by 1 or more behavioral constructs, the behavioral constructs themselves are not the results of the analysis.

RESULTS
We identified 6 behavioral barriers to and facilitators of MNP use that operate at 3 key points along the intention-to-behavior continuum: (1) accessing MNP through the health system, (2) forming an intention to use MNP, and (3) using MNP at the moment of child feeding. Examples of qualitative data illustrative of and supporting the behavioral constructs underlying each barrier/facilitator are available in Supplement 2.

Accessing MNP Through the Health System
1. Caregivers’ Experiences With Health Care Providers Shaped Their Motivation to Access MNP
Caregivers accessed MNP and received counseling about its use at health clinics during well child checkups. Because health care providers were seen as authority figures, the emotional valence of these interactions was influential. Some caregivers reported that negative interactions with health care providers (e.g., feeling dismissed, condescended to, rushed, shamed) made them reluctant to return to health establishments. Negative interactions may have also dissuaded caregivers from taking the advice about MNP given by that professional at the visit. Unclear or contradictory information from health care providers about MNP and its use could have led to ambiguity aversion. Some caregivers avoided using MNP if they felt they lacked sufficient information about it, especially if they were unsure about its potential harms. They expressed uncertainty about both effectiveness and administration of MNP. We originally posited that caregivers might also have avoided asking health care providers questions to resolve these doubts about MNP due to social desirability bias or wanting to appear competent in front of health care providers, but this was not supported during the validation step based on observations made during data collection.

In other circumstances, interactions with authority figures could have also acted as facilitators for MNP use. For example, sometimes it was easier for caregivers to give MNP if they felt like someone else (health care providers, family members) told them they had to give MNP, essentially making the decision for them. However, it should be noted that a few caregivers were uncomfortable with health care providers presenting MNP as a requirement, especially if they already doubted its quality or felt like they had not been given explanations for the reasons to use it. Additionally, caregivers who felt they had received good information from authority figures felt more confident about using MNP. For example, caregivers expressed satisfaction when health care providers took the time to explain MNP to them in depth and address their doubts. Finally, framing effects (how health care providers framed anemia to caregivers) affected their likelihood of administering MNP. For example, caregivers reported they were more likely to give MNP if they were told how anemia could affect their child’s brain and development, whereas being told their child had “low hemoglobin” was confusing and did not instill a sense of urgency.

2. Caregivers Felt Accessing MNP at Clinics Was Inconvenient and Created Hassle Factors
Even the smallest amount of friction or hassle reduced the probability that caregivers would access MNP or seek information on its use. If caregivers were required to attend informational sessions at inconvenient times or knew they would have to wait in long lines for a well child checkup (where they received MNP and counseling), they may have been reluctant to make the visit. We originally also posited that once caregivers already had MNP at home, they were more likely to give it to their child, an example of endowment effect, a psychological phenomenon where people are more likely to keep something they already have than make the effort to obtain it. However, this effect was not validated by other researchers given that caregivers were given an exact amount of MNP to last them between checkups (approximately 90 sachets) and thus would not have had any extra that would allow for prolonged use. Losing MNP sachets or other problems related to...
the number of sachets received was not a salient theme in our analysis.

**Forming an Intention to Use MNP**

3. Caregivers’ Mental Models About Anemia Prevention Shaped MNP Intentions and Use

Caregivers’ mental models about nutrition, how MNP works, and what could serve as a substitute for MNP directly shaped intentions to use MNP. Caregivers who perceived multiple ways to treat or prevent anemia, such as a variety of over-the-counter medications, may have experienced choice overload and may have looked for simplifying heuristics to choose among known alternatives, such as salience, familiarity, or ease. Some caregivers expressed the belief that over-the-counter medications were of higher quality than MNP, given that they were distributed by pharmacies (instead of the public health system) and had a monetary cost (as opposed to free distribution). Caregivers’ preferences to treat anemia through diet reflected the mental model (sustained primarily by peer or family advice but also recommended by health care providers) that a “natural” solution was preferable to a pharmacological one. Additionally, caregivers often defaulted to the use of food, a traditional and automatic response to treating illness, prompted by its presence in the home. A few caregivers expressed distrust of MNP because of its manufacture in India, which we hypothesized may have reflected a possible “not invented here” bias that may have limited regular use. During the validation process, other researchers concluded that this barrier may not have been as salient as other heuristics given its less frequent occurrence in the data. However, other mental models that drew analogies between MNP and something more familiar could have facilitated MNP use. Caregivers who described MNP as being “like vitamins” (versus medication) appeared more likely to have favorable views of MNP and to feel comfortable using it.

**Using MNP at the Moment of Child Feeding**

5. Caregivers Forgot to Give MNP if They Did Not Have Cues to Remind Them but Could Be Prompted With Salient Cues

Even when individuals intend to do something and have the resources to do it, they often require a specific prompt from the environment to overcome inertia. If the social, physical, or media environment failed to cue MNP use at the right time and in the right way, caregivers may have defaulted to nonuse. For example, some caregivers did not feel prompted to use MNP if their child appeared to be healthy, even if the child had been diagnosed with anemia. It may have been possible that they experienced “ostrich effect,” or unwillingness to accept this diagnosis for fear of dealing with the repercussions. The authors involved in data analysis originally posited that if caregivers did not see others in their communities using MNP, they may have assumed that others did not approve of MNP or simply that not seeing peers use MNP could have failed to cue MNP use. This was not supported during validation. Another early proposed barrier was that working or busy caregivers who bought prepared food instead of cooking homemade meals were less likely to add MNP to the purchased food because they did not go through the process of preparing and serving the food themselves. This, however, was eliminated during validation given that the relevant data referred to general nutrition practices, not MNP use. Alternatively, salient, well-timed cues from the environment could have promoted MNP use. Caregivers strongly suggested that additional information about MNP on
mass media, particularly television and radio, could have been an effective cue.

6. Caregivers Were Affected by Emotional, Cognitive, and Attentional Factors During Feeding That Were Difficult to Anticipate

The challenges of using MNP in the moment during child feeding could have led to procrastination and avoidance, exacerbated by the hot-cold empathy gap. When caregivers were at the clinic and decided to use MNP, they were in a deliberative and rational or “cold” state. This made it difficult to envision what it would have been like to apply MNP at mealtime, when caregivers were in an agitated, cognitively taxed “hot state” due to child experiences with side effects or dislike of taste. The counseling that caregivers received at well child checkups did not acknowledge this gap or help caregivers plan for it. In the validation process, other researchers confirmed that aversion to MNP at the moment of use led to its avoidance, especially if caregivers did not feel prepared to address any complications that may arise.

Some caregivers also over-focused on specific details of MNP administration, known as focusing effect, which made MNP easy to abandon if they felt they could not perfectly follow the instructions. Because skipping or incorrectly implementing a step may have resulted in worse taste or side effects (e.g., leaving it in food for an extended period of time increased the metallic taste due to capsule breakdown), this focus was understandable, but it may have led to abandoning MNP administration in the moment after a minor deviation from the protocol. Alternatively, when caregivers were removed from the process of administering MNP to the child, they expressed greater satisfaction with MNP and its effect on their child’s health. This was the case for caregivers who used Cuna Más, a public daycare that required them to bring in MNP with their child so the staff could administer it to the child during the day.

DISCUSSION

We identified 6 behavioral barriers to and facilitators of using MNP for anemia prevention among caregivers of young children in Arequipa, Peru. These are: (1) caregivers’ experiences with health care providers shaped their motivation to access MNP; (2) caregivers felt accessing MNP at clinics was inconvenient and created hassle factors; (3) caregivers’ mental models about anemia prevention shaped MNP intentions and use; (4) caregivers’ salient negative experiences could have caused them to stop giving MNP; (5) caregivers forgot to give MNP if they did not have cues to remind them, but could be prompted with salient cues; and (6) caregivers were affected by emotional, cognitive, and attentional factors during feeding that were difficult to anticipate.

Our results support and extend previous findings in the literature and provide opportunities for designing and revising program interventions that incorporate the behavioral constructs underlying the barriers.

First, as in other studies, we found that authority figures including health care providers have an opportunity to influence uptake and ongoing use of MNP.7,13,17,19 Our results place unique emphasis on the emotional context of these interactions. Caregivers had different perspectives on what constituted a positive or negative interaction with health care providers; some preferred an authoritative approach and others preferred collaborative decision making. Although research in Peru and many other countries promotes the use of culturally appropriate counseling techniques,13,17,26 research on collaborative decision making between health care providers and patients regarding nutrition is limited to the United States.37 Future research could examine whether authoritative or collaborative counseling styles would be most effective in motivating caregivers in Peru. In addition, as in other studies, we found that confusing information during consultations could have led to caregivers feeling they lacked sufficient information on MNP; this reinforces the need to simplify and tailor the educational campaigns recommended by other studies.13,19,26 Our results also highlight that framing anemia as “low hemoglobin” reduced both salience and urgency for caregivers; a higher salience framing in Ministry of Health campaign and health care provider training materials could emphasize children’s growth and brain development.

Second, our study highlights hassle factors as another major barrier when accessing MNP at clinics. Although prior studies have identified barriers to accessing health services, few connect those barriers to MNP adherence. Prior behavior science research has demonstrated that even minimal friction in a health or benefits program reduced take-up.46 In the context of MNP use, hassle factors reduced caregivers’ likelihood of accessing MNP and receiving information on its use. Therefore, our results point to the need for structural changes within the health system (more staff available for appointments, creating appointment and informational session schedules outside of caregivers’ work hours, and increasing access to MNP sachet “refills” in community settings). Cuna Más, a public daycare giving caregivers confusing information during visits could have led to them feeling they lacked sufficient information on MNP—reinforcing the need to simplify and tailor educational campaigns.
system in Peru, and community health promoters are trusted sources that could expand access to MNP.

Our third and fourth results concerned forming an intention to use MNP that we demonstrated was shaped by mental models about nutrition and by negative experiences with MNP. Caregivers may have had a preference for addressing anemia through diet, a more familiar and less “medicalized” approach than MNP. This default preference for dietary approaches may have led caregivers to attend to and implement dietary suggestions rather than use MNP even when health care providers made both recommendations. Our results also revealed that a way to frame MNP to align with the caregiver mental models about nutrition was as a vitamin supplement rather than as a medication. Negative experiences with MNP, such as side effects or bad taste, and the effect of negative comments from family and peers, have all been noted in previous studies. Although previous work has typically interpreted negative comments as lack of social support for MNP use and therefore proposed increased informational outreach to family members and peers as a solution, our results highlight the importance of including specific behavioral guidance to family and peers to not to overemphasize prior complications and negative experiences.

For any behavior that is new, challenging, and not yet habitual, it’s easy to procrastinate. In our fifth result, we confirmed a common finding from previous studies that caregivers needed external cues to overcome procrastination around MNP use at mealtime. General prompts in the form of television and radio spots were useful, and they have been used in Peru by Ministry of Health and other organizations to effectively promote anemia-specific and other positive health behaviors. Our approach also uncovered the importance of specific, timely, unavoidable cues at the moment of child feeding, which may drive behavior change more than a TV or radio spot heard earlier in the day. Possible interventions informed by this insight include encouraging caregivers to store MNP sachets with other items that will be used during mealtime (i.e., with the child’s dish or utensils) or sending an SMS message reminder at common mealtimes.

Finally, we identified “hot-cold empathy gap” as a barrier to consistent MNP use. Hot-cold empathy gap has been observed in other health behaviors, in which people consistently fail to imagine and account for what a future “hot” affective or cognitive state will be when a plan to act is made ahead of time in a “cold” state. Hot-cold empathy gap is particularly relevant for understanding what happened when caregivers intended or planned to use MNP during child feeding but abandoned that plan in the moment when children resisted, refused, reacted strongly to taste, or experienced side effects such as diarrhea, all of which are commonly cited barriers in the literature. Awareness of the hot-cold empathy gap can provide a channel for improved program development. Given that people generally don’t understand their actions as “state-dependent,” a possible intervention involves preparing caregivers for side effects that might cause anxiety. Health care professionals could encourage caregivers to think through how they would react to a stressful experience while in a “cold state” at their medical appointment. This strategy is based on evidence from the side effect reduction literature, often focused on cancer patients, that suggests that preparation for side effects can reduce anticipatory symptoms and stress and improve coping skills. This approach would represent a departure from current Peruvian Ministry of Health trainings that teach health care providers to counsel caregivers that MNP has no side effects.

Limitations

There are several limitations to this study. Brewer et al. describes the limitations in study design and data collection, such as recruitment of participants in and around health centers, the reliance on self-reported (rather than observed) barriers to MNP use, and the lack of systematic collection of sociodemographic and behavioral characteristics of the caregivers and their children (such as time using MNP, birth order, etc.). Adopting a behavioral economics perspective limits the identified barriers and facilitators to those with a specific behavioral (as opposed to structural) underpinning. Additionally, the behavioral barriers to optimal MNP use identified in this analysis require further confirmation through empirical testing of interventions designed to address them. Although NUDGE is similar to several other approaches using behavioral economics and human-centered design, it is still evolving as an analytic tool and future iterations may further refine the approach.

CONCLUSION

This study uses behavioral economics and a behavioral design approach to understand MNP administration and childhood anemia prevention generally. This approach to analyzing cognitive...
biases and heuristics can generate insights into behavioral influences on adherence that complement existing approaches to identifying barriers to take-up of evidence-based practices. Our results led us to focus on various underlying heuristics that influenced MNP adherence, such as authority bias and framing effects, hassle factors, the salience of negative experiences and certain mental models, appropriately timed cues, and the hot-cold empathy gap. Consideration of these behaviors and underlying biases may inform aspects of programmatic intervention, such as counseling practices to promote MNP use, to improve adherence in an area of Peru that has a disproportionately high burden of anemia.

Acknowledgments: We would like to thank Dr. Ricardo Castillo-Neyra, Dr. Michael Z. Levy, Katty Borrini Mayorí, Lina M. Mollecoa Riveros, Amparo M. Toledo Vizcarra, and the entire field team at the Zoonotic Disease Research Lab in Arequipa, Peru, for their role in project implementation and data collection. We would also like to acknowledge Dr. Richard A. Oberhelman for his role in securing funding in the initial data collection for this project and providing research guidance.

Funding: This work was supported by the National Institute on Minority Health and Health Disparities, “Tulane-Xavier Minority Training in International Health” through the Minority Health International Research Training Program [grant number T37 MD001424]; the Stone Center for Latin American Studies at Tulane University with the Tinker Foundation; and the Fulbright U.S. Student Program. The conclusions expressed in this article are entirely those of the listed authors and are not endorsed by the Fulbright Program, the U.S. Department of State, or any of its partner organizations. None of the funders had any role in study design, data collection and analysis, decision to publish, or manuscript preparation.

Competing interests: None declared.

Author contributions: KR and VPS were involved in study design and administration. Data was collected by JDB, JS, KR, and MPS. Analysis was conducted by JDB, JS, AB, KR, and MPS. JDB, JS, and AB prepared the manuscript, and KR, MPS, and VPS provided input and edited various drafts.

REFERENCES
Medline


La anemia, aplicamos un enfoque del diseño del comportamiento (behavioral design). La anemia infantil sigue siendo un factor importante de morbilidad en los países de ingresos bajos y medios, incluyendo al Perú. Para identificar los retos conductuales para el uso de los micronutrientes en polvo (MNP) y una estrategia para enfrentar los retos que podrían encontrar en el uso de los MNP.

Implicaciones claves:

- Las interacciones entre el personal de salud y los cuidadores de niños y sus estados emocionales cambiantes, desde que forman la intención para usar los micronutrientes en polvo (MNP) hasta que los usan afectaron su adherencia a los MNP, proporcionados por el gobierno para prevenir la anemia infantil. Durante la consejería para cuidadores, sugerimos que el personal de salud proporcione mensajes claros sobre el impacto de los MNP y una estrategia para enfrentar los retos que podrían encontrar en el uso de los MNP.

Hallazgos claves:

- Experiencias negativas con el personal de salud o inconveniencias al momento que se inicia el uso de los MNP desalientan el uso futuro de los MNP.
- Modelos mentales sobre la nutrición pueden influir en las intenciones para usar los MNP y, a la vez, el hecho de tener demasiadas opciones puede confundir a los cuidadores.
- Una sola experiencia negativa con los MNP puede formar memorias impactantes y desalientar a los cuidadores a dar los MNP.

Implicaciones claves:

- La capacitación para el personal de salud debería fomentar una interacción interpersonal positiva con los cuidadores durante la iniciación de los MNP para que estas interacciones puedan tener un impacto duradero en el uso de los MNP.
- La consejería para cuidadores debería incluir información sobre efectos secundarios que podrían tener los niños para estar preparados para manejarlos.
- En programación futura, los profesionales de la salud pública deben considerar alentar a los cuidadores a utilizar señales oportunas para administrar MNP.

Resumen

La anemia infantil sigue siendo un factor importante de morbilidad en los países de ingresos bajos y medios, incluyendo el Perú. Para identificar los retos conductuales para el uso de los micronutrientes en polvo (MNP), administrados con el fin de complementar la alimentación de los niños y prevenir la anemia, aplicamos un enfoque del diseño del comportamiento (behavioral design) en entrevistas y grupos focales realizadas con 129 cuidadores en Arequipa, Perú. Examinamos 3 puntos claves en el proceso de toma de decisiones: acceso a los MNP a través del sistema de salud; formación de la intención para usar los MNP; y el uso de los MNP al momento de alimentar al niño. Utilizando la estrategia de NUDGE (de las siglas en inglés para Reducir, Entender, Descubrir, Generar y Evaluar), identificamos las siguientes barreras y facilitadoras de comportamiento. (1) Las experiencias de los cuidadores con el personal de salud determinaron su motivación para acceder a los MNP (2) Los cuidadores sintieron que acceder a los MNP en las clínicas era inconveniente y creaba molestias; (3) Los modelos mentales de los cuidadores sobre la prevención de anemia influyeron en la intención de dar los MNP; (4) Una experiencia negativa marcada de los cuidadores pudo haber hecho que dejaran de dar los MNP; (5) Los cuidadores se olvidaban de dar los MNP si no tenían señales (cues) para recordárselos, pero podían recibir señales oportunas para su administración; y (6) Los cuidadores fueron afectados durante la alimentación por factores emocionales, cognitivos y de atención, que fueron difíciles de anticipar. Nuestros resultados, basados en un enfoque del diseño del comportamiento, sugieren oportunidades de adaptar la comunicación, orientación y educación actual sobre el uso de los MNP. Las modificaciones incluyen proporcionar una comunicación culturalmente relevante, aprovechando los estados emocionales y cognitivos de los cuidadores, y alentar cambios pequeños pero impactantes en la rutina de alimentación para combatir las barreras al uso de los MNP.

Perspectivas Conductuales Sobre el Uso de Micronutrientes en Polvo Para la Anemia Infantil en Arequipa, Perú

Mensaje clave: Las interacciones entre el personal de salud y los cuidadores de niños y sus estados emocionales cambiantes, desde que forman la intención para usar los micronutrientes en polvo (MNP) hasta que los usan afectaron su adherencia a los MNP, proporcionados por el gobierno para prevenir la anemia infantil. Durante la consejería para cuidadores, sugerimos que el personal de salud proporcione mensajes claros sobre el impacto de los MNP y una estrategia para enfrentar los retos que podrían encontrar en el uso de los MNP.

Hallazgos claves:

- Experiencias negativas con el personal de salud o inconveniencias al momento que se inicia el uso de los MNP desalientan el uso futuro de los MNP.
- Modelos mentales sobre la nutrición pueden influir en las intenciones para usar los MNP y, a la vez, el hecho de tener demasiadas opciones puede confundir a los cuidadores.
- Una sola experiencia negativa con los MNP puede formar memorias impactantes y desalientar a los cuidadores a dar los MNP.

Implicaciones claves:

- La capacitación para el personal de salud debería fomentar una interacción interpersonal positiva con los cuidadores durante la iniciación de los MNP para que estas interacciones puedan tener un impacto duradero en el uso de los MNP.
- La consejería para cuidadores debería incluir información sobre efectos secundarios que podrían tener los niños para estar preparados para manejarlos.
- En programación futura, los profesionales de la salud pública deben considerar alentar a los cuidadores a utilizar señales oportunas para administrar MNP.
A Cluster-Randomized Trial to Test Sharing Histories as a Training Method for Community Health Workers in Peru

Laura C. Altobelli, a,b,c José Cabrejos-Pita, b,d Mary Penny, e Stan Becker f

Key Findings

- Sharing Histories method for training community health workers (CHWs) was associated with reduced child stunting compared with a standard CHW training method.
- The training method’s impact on child stunting was not present when mothers were illiterate.

Key Implications

- CHWs can learn better by using Sharing Histories as the basis for their training; then, they can use the same method to better help mothers change behaviors.
- Program managers can use this training methodology to strengthen capacities of CHW trainers to improve the cultural literacy of trainees.
- Primary health care personnel who train CHWs can identify local cultural and social norms when CHWs share their experiences of childbearing and childrearing.
- Primary health care personnel can easily apply this CHW training method to develop social bonds with CHWs, address CHW cultural competencies, and provide CHWs with a strategy to effectively discuss sensitive culturally determined behaviors with mothers.

ABSTRACT

Background: Community health workers (CHWs) are increasingly deployed to support mothers’ adoption of healthy home practices in low- and middle-income countries. However, little is known regarding how best to train them for the capabilities and cultural competencies needed to support maternal health behavior change. We tested a CHW training method, Sharing Histories (SH), in which CHWs recount their own childbearing and childrearing experiences on which to build new learning.

Methods: We conducted an embedded cluster-randomized trial in rural Peru in 18 matched clusters. Each cluster was a primary health facility catchment area. Government health staff trained female CHWs using SH (experimental clusters) or standard training methods (control clusters). All other training and system-strengthening interventions were equal between study arms. All CHWs conducted home visits with pregnant women and children aged 0–23 months to teach, monitor health practices and danger signs, and refer. The primary outcome was height-for-age (HAZ)< −2 Z-scores (stunting) in children aged 0–23 months. Household surveys were conducted at baseline (606 cases) and 4-year follow-up (606 cases).

Results: Maternal and child characteristics were similar in both study arms at baseline and follow-up. Difference-in-differences analysis showed mean HAZ changes were not significantly different in experimental versus control clusters from baseline to endline (P=.469). However, in the subgroup of literate mothers, mean HAZ improved by 1.03 on the Z-score scale in experimental clusters compared to control clusters from baseline to endline (P=.059). Using generalized estimating equations, we demonstrated that stunting in children of mothers who were literate was significantly reduced (Beta=0.77; 95% confidence interval=0.23, 1.31; P<.01), adjusting for covariates.

Conclusion: Compared with standard training methods, SH may have improved the effectiveness of CHWs as change agents among literate mothers to reduce child stunting. Stunting experienced by the children of illiterate mothers may have involved unaddressed determinants of stunting.

INTRODUCTION

Supporting mothers to adopt healthy home practices could be one of the keys to improving child health. Health behavior change strategies are frequently used in global health, but the evidence of their effect on health
CHWs are deployed worldwide, but their effective outcomes remains unclear due to numerous pitfalls in their design and evaluation. A central challenge for these strategies is how to effectively help mothers gain knowledge and change behaviors in communities with strong traditional beliefs and poor access to health information. A worldwide priority for child health and development is the reduction of chronic child malnutrition (poor linear growth or stunting), which arises from a broad range of causes related to home practices for maternal nutrition, breastfeeding and weaning, water, sanitation, hygiene, and infection prevention. Peru is a low- to middle-income country (LMIC) that has seen a major overall reduction in stunting since 2008, but high rates persist in mountain and jungle regions where numerous cultural practices negatively influence maternal and child health.

Community health workers (CHWs) are a global priority to help reach impact and equity goals through universal health coverage and Sustainable Development Goals. CHWs are the lowest level of frontline health workers and are frequently volunteers, delivering a wide range of services in homes and communities including health education and support on nutrition, malaria, tuberculosis, HIV/AIDS, sexually transmitted infections and noncommunicable disease, preventive maternal and reproductive health services in the home, management of uncomplicated childhood illnesses, and access to services, among others. CHWs can be critical actors for reporting maternal and perinatal deaths occurring in the community. We know some of what works, but a large gap remains between that knowledge and “how to make it work.” An estimated 5 million CHWs are deployed worldwide, but their effectiveness and linkage to health subsystems within their communities vary. The World Health Organization (WHO) identifies CHWs as important to their Global Strategy for Human Resources, but implementation research on CHW programs is needed.

To be effective, CHW programs should have detailed plans for governance/management, selection, training, supervision, engagement with communities, relationship with the health system, scaling up, and monitoring and evaluation. In this study, we highlight the need for the identification and testing of the best methods to train CHWs. We have not found other reports with results on comparative studies. We submit that special training methods are needed to adequately prepare CHWs to support home behavior change. Behavior change theories developed for industrialized countries often cannot be applied in areas of LMICs with embedded cultural beliefs, attitudes, and practices. More recently, integrated behavior change models for LMICs have been developed using a mix of theories and strategies. Even though qualitative research can identify the specific beliefs that impede the home practice of key health behaviors, defining the “black box” of how and why mothers, families, and communities hold onto cultural beliefs is a major challenge when working to empower change agents who can convince mothers to change cultural practices.

An important part of empowering CHWs is helping them build their own self-confidence and agency so they can thus empower mothers. These are key dimensions of maternal capabilities needed to implement new knowledge of proper child care. We assume that CHW efforts empower communities, but it is less commonly recognized that CHWs themselves need to become empowered to be the change agents needed to support mothers and families for active self-care. According to Kane and colleagues, “...to be able to empower the communities they serve, we argue, it is essential that CHWs themselves be, and feel, empowered....” A review of randomized controlled trials with a “realist” approach concluded that interventions by CHWs worked if there was a “...sense of relatedness with beneficiaries and public services; increase in self-esteem; sense of self-efficacy....” The same author concluded that if these factors were absent, CHW performance would be negatively affected even with the same interventions. Some researchers have suggested the existence of a “secret sauce” that would help to empower women with knowledge, motivation, and increased self-efficacy even when scaling up community strategies into government programs. This “sauce” could be the next breakthrough to sustainably improve maternal, newborn, and child health behaviors.

How to maintain fidelity of empowerment and behavior change approaches in the scale-up of interventions with CHWs remains a key challenge. As a part of being and feeling empowered, the cultural competency of health workers and CHWs is an essential skill to reach patients of diverse cultures to improve their health literacy. Scaling up CHW programs in government systems likely relies on how well health providers in primary health care (PHC) services can serve as trainers to facilitate CHW learning. They are generally not educators, and they often rely on medical terminology and heuristic methods to train CHWs. Methods for teaching CHWs should be adapted to their educational level, which is very often the same as that of the mothers with whom they will work. The train-the-trainer model, CHW training materials, CHWs themselves need to become empowered to be the change agents needed to support mothers and families for active self-care.
incentives, and supportive supervision are factors that have yet to be refined in government systems as well as nongovernmental organization and other private sector efforts to support CHWs.

This study provides a new focus in the development of empowerment and cultural competency by testing a training method that builds on cultural beliefs and practices of CHWs so they increase their self-esteem and have a greater sense of interpersonal relatedness with households and the health system. From this foundation, they can become more effective behavior change agents within traditional populations.

We conducted a cluster-randomized controlled trial of our new teaching strategy called Sharing Histories (SH) in rural Peru to test the impact on child stunting when female CHWs are trained with this method. We hypothesized that mothers would be more likely to change health knowledge and behaviors, and their children would consequently have better growth, if the mothers received health information from CHWs trained with the SH method, compared with the situation in which health education interventions were received from CHWs trained with standard methods.

The method is used to train CHWs and community supervisors (CSs), as well as to provide direct education and counseling of mothers, and it can be implemented at low cost through the government PHC system. The training method is oriented toward enhancing CHWs’ empowerment, as well as their knowledge and skills, as community change agents.

### SHARING HISTORIES TRAINING METHOD

The SH teaching method is a key component of an integrated teaching strategy we developed to improve maternal, neonatal, and child health. It was piloted in Afghanistan, India, and Peru.\textsuperscript{35–38}

Research on autobiographical memory explains how the Sharing History method helps CHWs learn to effectively convince other women to change behaviors.\textsuperscript{38,39} The utility of memory fits into 3 categories: self function (continuity of the self), social function (developing and maintaining social bonds), and directive function (guiding present and future behaviors).\textsuperscript{40,41} The latter 2 functions are relevant to SH. The social function creates bonding and trust among CHWs who share their childbirth and childrearing memories with each other, and these memories are then used by trainers as the basis for learning. The directive function has been found to influence future behavior.\textsuperscript{40} Neuroscience research suggests that physical consolidation of neurons occurs when memory is stimulated and new information is added.\textsuperscript{42}

A manual provides details of the SH training method and the community health model (Box).\textsuperscript{43}

CHW trainees can take ownership of and learn from their own cultural beliefs and practices, which empowers them to change their practices and helps them become more capable and convincing promoters of the same key behavioral changes with mothers in their communities. The process of sharing personal experiences could have several benefits including providing CHWs, who may be timid about speaking in a group, with opportunities to practice verbal expression and revealing cultural

---

**The Sharing Histories training method is oriented toward enhancing CHWs’ empowerment, as well as their knowledge and skills, as community change agents.**

**BOX. Teaching Strategy of Sharing Histories**

The teaching strategy of SH builds on CHWs sharing their personal experiences and actions regarding their pregnancies, births, postpartum periods, care and feeding of newborns and infants, and events surrounding any sickness or death. Monthly full-day workshops are held in the primary health care facility. Each of the 7 training module topics has a series of class sessions with 6 steps:

1. If pregnancy is the day’s topic, each CHW shares her pregnancy experiences while the trainers or assistants take notes on the history format.
2. Trainers list key actions mentioned by CHWs, then lead CHWs through a guided discussion of each action, using colors to indicate whether the action is beneficial, neutral, or potentially harmful and then discussing why.
3. Then, picture cards on key practices are used to teach each best practice in a class session, referring to the CHWs’s shared experiences and further asking about and analyzing local customs related to each best practice.
4. Participatory methods are then used to practice what is learned in each class session, such as sociodramas of home visits to teach mothers using the same methods and materials, monitor practices and danger signs, and make referrals.
5. Each class session ends with participatory evaluation of learning, using games and exercises.

Between monthly class sessions, each CS meets with 5 or 6 of her assigned CHWs to review the monthly topic, and CHWs practice using the respective flip charts and checklists for monitoring key maternal health practices and danger signs.
practices that can be used to build new learning that would otherwise only be identified through expensive and time-consuming qualitative research.

**METHODS**

**Project Intervention**

This study was an embedded component of a larger integrated project intervention called Health in the Hands of Women, which was aimed at reducing chronic child malnutrition by linking strengthened PHC services and district government to a sustainable community-oriented system to support maternal behavior change for improved maternal, neonatal, and child health. The area where this study was conducted included 3 rural districts with 82,000 inhabitants in the area of the upper Huallaga River on the eastern slope of the Andes mountains in the Huánuco Region of Peru (Figure 1). In both experimental and control study arms, we implemented interventions to strengthen capabilities and processes so that district government, health services, and local health administration committees could better support CHWs. The methodology of CHW training was the only factor that varied between the 2 study arms (Table 1).

As part of the larger project, selection criteria, roles, and tasks were clearly defined for each type of human resource for community health: CHWs, CSs, and PHC facility staff who train CHWs and CSs (Table 2).

Key project messages for training CHWs and CSs and for teaching mothers were identified from best-practice literature on reducing child stunting and from our baseline qualitative studies on local practices of the target population. Key messages were delineated in a series of 7 flip charts focused on 7 areas of maternal knowledge and practice during the first 1,000 days from conception. These 7 flip charts covered the topics: pregnancy, birth and postpartum, newborns, breastfeeding, child growth and nutrition, infant diarrhea, and infant pneumonia. Each flip chart emphasized home practices, preventive care services, and recognition...
of danger signs for which medical care should be sought. Key messages on water, sanitation, and hygiene (WASH) practices were included throughout the 7 flip charts. The flip chart messages and artwork were previously developed and validated by the research team in another rural area of Peru (Cusco)46 with artwork adapted to reflect local clothing and hair styles of Huánuco. The breastfeeding flip chart was adapted from one previously developed in Lima.47

All trainers, CSs, and CHWs received the flip charts, which were identical for both study arms. For the trainers, 2 sets of step-by-step training manuals were developed so they could apply different methods for training CHWs on the flip chart messages: (1) a set that incorporated the SH teaching method and (2) a set that used standard participatory CHW training methods. CHWs in both study groups received and studied the same flip charts and used them in home visits to teach mothers.

A set of 12 checklists and reporting formats were other key tools taught to and used by CHWs equally in both study arms. These included individual pictorial checklists developed by the project team for home-monitoring of pregnant/postpartum mothers and infants aged 0–23 months46 and for newborns (checklist adapted from the SEARCH Program48). Other tools previously developed or adapted and validated were for community referral, supervision and reporting, mothers’ birthing plan to keep at home,46 and community development planning.49

Trainers, CSs, and CHWs were selected and trained between 2010 and 2014 in either the experimental or control teaching method based on their corresponding PHC facility cluster. Twenty-three selected health staff from 11 experimental PHC facilities were trained as trainers in the SH teaching methodology, and 23 staff from 11 control PHC facilities were trained separately in standard CHW training methods. Trainers received 14 nonconsecutive days of training: 6 days in adult learning methodologies and 8 days in use of the corresponding set of 8 training manuals by type of teaching methodology (1 for each of the 7 topics plus 1 introductory manual).

### TABLE 1. Comparison of CHW Training Methodology Interventions Used in a Cluster-Randomized Controlled Trial, Huánuco Region, Peru

<table>
<thead>
<tr>
<th>Interventions</th>
<th>Experimental Clusters</th>
<th>Control Clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic strengthening of primary health care services: orientation to community health strategies, interdisciplinary team building for health staff, self-assessment, and planning for community health actions</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Basic strengthening of local government to support community MNCH, to gain their commitment to provide financial and incentive support to CHWs and CSs</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Training of facilitators (health personnel trainers) on adult education methods</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Provision of a complete set of 7 flip charts to each CHW and CS</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Training of facilitators on use of facilitator manuals based on Sharing Histories as the CHW training method</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Training of facilitators on use of facilitator manuals based on a standard CHW training method</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Continuous monthly training of CHWs and CSs using Sharing Histories as the training method</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Continuous monthly training of CHWs and CSs using a standard CHW training method</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Monthly home visits conducted by CHWs, supported by CSs and health staff to educate mothers, monitor MNCH behaviors, identify danger signs, and refer to the health facility</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Monthly supervision of CHWs by CSs</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Abbreviations: CHW, community health workers; CS, community supervisors; MNCH, maternal, neonatal, and child health.
<table>
<thead>
<tr>
<th>CHW</th>
<th>CS</th>
<th>Facilitators (Trainers)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Selection process</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selected preferably by other women in the community</td>
<td>Selected by a panel of judges from the local PHC facility and municipal government</td>
<td>Self-selected with the approval of their superior</td>
</tr>
<tr>
<td><strong>Selection criteria</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respected older woman with grown children</td>
<td>Female, literate, at least 3 years of prior CHW experience or auxiliary nurse training; ability to work half-time and travel between communities</td>
<td>Health professional; preferably woman who has a long-term contract in PHC facility</td>
</tr>
<tr>
<td><strong>Workload</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 CHW for every 30 families (on average, 2 or 3 pregnant women and 2 or 3 children aged 0–23 months)</td>
<td>1 CS supports 10–15 female CHWs</td>
<td>1 trainer per group of 10–25 CHWs and their respective CS</td>
</tr>
<tr>
<td><strong>Key roles</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Attend monthly 1-day trainings at nearest PHC facility</td>
<td>• Ensure that her assigned CHWs attend month training sessions in the PHC facility</td>
<td>• Organize and hold 1 monthly workshop for CHWs in their own PHC facility</td>
</tr>
<tr>
<td>• Meet monthly in small groups with her CS for reinforcement of training and practice with flip charts and monitoring formats</td>
<td>• Meet with her CHWs in small groups of 5–7 CHWs once or twice a month to review the training from the latest workshop in the PHC facility and to practice using flip charts to teach mothers</td>
<td>• Train CHWs and CSs on how to educate and monitor mothers in the home using the flip chart series and monitoring tools, following a facilitator manual corresponding to each flip chart</td>
</tr>
<tr>
<td>• Create a map of her 30 households, identifying pregnant women and children aged 0–23 months</td>
<td>• Accompany CHWs on home visits until the CHW feels comfortable visiting alone</td>
<td>• Receive training in how to utilize the facilitator manual that accompanies each of 7 flip charts</td>
</tr>
<tr>
<td>• Visit each pregnant woman and child aged 0–23 months on a monthly basis</td>
<td>• Attend monthly training workshops along with CHWs in the PHC facility</td>
<td>• Use the Sharing Histories teaching methodology as incorporated into each facilitator manual</td>
</tr>
<tr>
<td>• Fill out simple monitoring checklists, referral slips, and monthly activity report checklists</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Key tasks during monthly home visits to pregnant women and children aged 0–23 months</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Share histories and teach mothers using flip charts by stage of pregnancy or child age</td>
<td>• Monitor health practices and record on pictorial checklists by stage of pregnancy or child age</td>
<td></td>
</tr>
<tr>
<td>• Monitor health practices and record on pictorial checklists by stage of pregnancy or child age</td>
<td>• Observe for danger signs using pictorial checklists by stage of pregnancy or child age</td>
<td></td>
</tr>
<tr>
<td>• Make referrals using pictorial referral slips for maternal-child preventive care and when danger signs are detected</td>
<td>• Make referrals using pictorial referral slips for maternal-child preventive care and when danger signs are detected</td>
<td></td>
</tr>
<tr>
<td><strong>Incentives</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-kind from the health system: certificate of recognition, training, supervision visits by CSs</td>
<td>Monthly stipend equivalent to about one-third the salary of an auxiliary nurse (To ensure accountability, monthly payment from district government was based on demonstrated completion of the 4 key roles)</td>
<td>Training and recognition</td>
</tr>
<tr>
<td>In-kind from district government: clothing items identifying her as a CHW or CS with name of the district, a food basket and party for annual Health Promoters’ Day and Christmas</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CHW, community health workers; CS, community supervisors; PHC, primary health care.
Monthly 1-day training workshops for about 500 CHWs and 46 CSs were organized and held by trainers in PHC facilities, unless roads were impassable during the January to March rainy season. Each of the 7 topics was taught in 1 to 3 full-day workshops. Each time the 7 topics were completed, another round of monthly workshops was initiated.

Study Objective
The study objective was to test the attributable effect of SH on child stunting: height-for-age less than -2 Z-scores below the median according to the WHO growth standard. The 2 study groups were defined as (1) mothers and children living in PHC facility catchment areas where CHWs were trained with the SH teaching method (experimental clusters); and (2) mothers and children living in catchment areas where CHWs were trained using standard methods (control clusters). Both study groups received home visits by experimentally trained or standard method trained CHWs, respectively.

Study Design
We conducted a cluster-randomized controlled trial (cRCT) in 22 clusters, with each cluster comprising a PHC facility and its catchment area population. A cRCT design was appropriate in this study based on criteria to select best methods to evaluate behavior change techniques. The cRCT design enabled overcoming the difficulties of using distinct training methods for individually randomized CHWs and potential contamination between study groups. This trial is registered with ClinicalTrials.gov, NCT02903602.

All PHC facilities and their catchment areas in 3 municipal districts were included. Clusters were matched in pairs on 2 criteria: category of PHC facility resolutive capacity and distance from district capital. Matched pairs were then randomly allocated to either study arm by the principal investigator. Interventions were applied at the cluster level for health personnel trainers, CHWs, and CSs. Type of CHW training method depended on the cluster to which they had been randomized.

At endline, 2 small communities did not fall into the systematic random sampling of sampling clusters (as distinguished from the randomized intervention clusters), due to the small size of those communities and the low number of children aged 0–23 months. Thus, it was necessary to exclude each of their respectively matched communities with which they had been matched before the baseline survey. As a result, 18 clusters were included in final data analyses. Outcomes were measured in 2 independent samples of households at baseline and endline. The allocation of study clusters is illustrated in Figure 2.

Blinding
The study was triple-blinded (participants, program implementers, and outcomes assessor). Ministry of Health and municipal officials were unaware of the 2 different teaching methods being used. PHC staff who served as CHW trainers were trained in separately programmed workshops and
were unaware of differences in the training methods being used. Experimentally trained CHWs and CSs had no contact with peers from control clusters. Household survey teams and data managers were blinded to the study groups.

**Project Evaluation**

Several data collection methods were used to assess baseline, process, and impact. Baseline qualitative studies were conducted on key home practices for maternal and child health and nutrition; CHW efforts in communities; views of community and municipal leaders of health actions; and assessments of PHC facilities in the project area regarding their level of organization for quality services and for work in communities. Project monitoring provided data on CHW training attendance, pre- and posttests of CHW learning in training workshops, and completion of tasks by CSs. At the project’s end, a sample of CHWs and all CSs were interviewed with close-ended questions. An in-depth qualitative study was conducted at endline with a sample of CHWs, CSs, and trainers. Most importantly, repeat household interview surveys were conducted at baseline and endline to evaluate the impact of the experimental training intervention.

**Methods for Household Interview Surveys**

Cross-sectional household surveys were conducted at baseline and at 4-year follow-up on 2 independent samples of mothers with children aged 0–23 months in the 2 study areas. The surveys were conducted, under guidance of the study team, by an independent research institution (Instituto de Investigación Nutricional) that was blinded to study groups.

**Study Variables**

The main outcome variable was stunting, height-for-age less than −2 Z-scores, as a proxy measure of health status in children aged 0–23 months. Independent variables previously shown to be associated with child stunting were measured, considering the following constructs: (a) birth weight, (b) breastfeeding practices, (c) child morbidity, (d) early home treatment of child illness, (e) use of health services for prevention and illness, (f) complementary child feeding practices, (g) micronutrient consumption in pregnancy and infancy, (h)
WASH practices, and (i) mothers’ knowledge of danger signs during the first 1,000 days. We also asked mothers about their receipt of benefits from selected government health and social services and any CHW home visits during and after their most recent birth.

**Survey Sampling Frame**
We obtained a list of children in each community from registers at each PHC facility. The study area had 22 health facility clusters with approximately 1500 children aged 0–23 months, yielding a sample that represented about 40% of these children.

**Survey Sample Size**
Given a baseline value of 35% stunting in children aged 0–23 months, we expected a reduction from 35% to 22% in the experimental study arm and from 35% to 30% in the control arm. With the number of children of eligible age in the study area limited to 1500, we applied a finite population correction to our sample size calculation.52 With a 95% confidence level (CI) and 80% power, each study group was estimated at 283 cases. An assumed 7% nonresponse rate gave a total of 303 cases per study group.

**Survey Sample Selection**
Two-stage sampling of households was conducted independently for experimental and control study arms. In the first stage for each study group, a list of communities and children aged 0–23 months in each randomized study cluster (PHC facility catchment area) was obtained from PHC facility registries. From this list, 38 sampling clusters (as distinguished from study cluster) were randomly selected by a systematic process with a random start. The number of sampling clusters in a community was proportional to its size, and each cluster had 8 children.

The second stage of sampling for household surveys was conducted upon arrival in a selected community. Permission to interview was obtained from village leaders who also helped to identify children born in the previous 4 months, who were then added to the list from the respective health facility registry. Children’s names were alphabetically listed and systematically selected with a random start. This method avoided selecting 2 children from the same family. If a community had fewer than the required number of eligible children, the team moved to the next closest community to reach the needed number.

**Survey Instrument**
The survey questionnaire was adapted from the Demographic and Health Survey instrument to allow direct comparison of results with national and international survey data.3 Questions were added to measure exposure to CHW visits and teaching materials, as well as empowerment and WASH indicators.

**Survey Team**
Experienced local field interviewers and anthropometrists were competitively selected and trained for baseline and final surveys. One-third of each team had full or working knowledge of Quechua, the indigenous language. Anthropometrists (nurses) received standardized training for child length and weight measurement and were assisted in the field by the interviewer. Training of survey teams lasted 1 week and focused on the consent process and use of the instrument, with supervised practice with mothers and their young children from outside the study area.

**Anthropometry Measures**
Digital platform scales were used to measure child weights, and their precision and accuracy were checked to 100 g before use. The mother was weighed alone first and then with the child wearing only light clothing. Both weights were recorded for later consistency checking. Length was measured using a lightweight folding durable plastic infant-o-meter accurate to 1 mm,53 with a minor modification to prevent movement of the foot board if the child pushed against it.

**Survey Field Supervision**
Supervisors accompanied each fieldworker to check survey forms in the field for completeness, and to conduct periodic repeat surveys with 10 questions after the main interview by the fieldworker.

**Survey Data Entry and Analysis**
Double data entry was done using Visual Fox Pro. Consistency and range checks were also done, and data were checked against the original forms as needed. Breastfeeding and infant feeding practices based on 24-hour dietary recall were evaluated using standard WHO indicators.54 Child anthropometry measures were converted to Z-scores of height-for-age, weight-for-age, and weight-for-height, using the 2006 WHO growth standard.50

We conducted an intention-to-treat analysis considering all sampled subjects in the study areas. Univariate assessment and bivariate tests of association were conducted on a wide range of
independent variables to establish comparability of study groups and to identify potential predictors of child stunting. Bivariate regressions of the outcome stunting on independent variables were assessed with generalized estimating equations (GEE) to adjust for clustering.

We conducted 2 stages of a difference-in-differences (DID) analysis to test the differences from baseline to endline of the experimental intervention on children’s mean height-for-age Z-score (HAZ). We first conducted a standard DID analysis for all children, then by subgroups stratified by low and high maternal literacy. The DID analyses compared mean HAZ by 9 pairs of matched clusters, comparing baseline to endline. Next, we compared 2 levels of maternal literacy within each study group to quantify the effect of the observed interaction of the experimental intervention by stratified levels of high and low maternal literacy on stunting.

Finally, we built a GEE model on endline data to determine the effect of the experimental intervention in interaction with maternal literacy versus the effect of the control intervention on the outcome stunting. Covariates with a \( P \)-value of .20 or less in the bivariate regressions on stunting were tested for inclusion in the multivariate model. Covariates that were colinear with the outcome variable were not included. Data were analyzed using SPSS version 17 for the baseline and version 20 for the final survey.

**Ethics and Informed Consent**

The Institutional Research Board of the Instituto de Investigación Nutricional approved the household survey proposal and consent process. Consent was verbal using an approved standardized protocol.

**Methods for Close-Ended Interviews With CHWs and CSs**

Fifty CHWs from each study group (n=100) were interviewed after the intervention period to identify their sociodemographic characteristics and perceptions of the training received, their roles in teaching mothers, their participation in CHW groups with their CS, and changes in themselves as a result of the training. CHWs were randomly selected from a list of active CHWs in each study cluster, proportional to the number of CHWs in each cluster. Interviews were conducted by nursing students from the local university following training and practice in interviewing techniques from expert interviewers. We used a closed-ended questionnaire adapted from Care Group program materials.\(^{55}\)

All CSs (46) were interviewed following the intervention period for the same reasons as listed for the CHW interviews. In addition, the CS interviews aimed to identify their perceptions of collaboration with health personnel, community leaders, and local government. We adapted a closed-ended questionnaire developed for Care Group programs.\(^{39}\)

**Methods for In-Depth Qualitative Study**

Following the intervention period, to triangulate quantitative findings, individual in-depth interviews were conducted with trainers, CSs, and CHWs from the experimental group on their experiences and opinions regarding SH as a teaching/learning methodology. Control group participants were interviewed on their training and learning experiences. Informed consent was obtained for all interviews. Using a unique interview guide for each type of respondent, interviews were recorded, translated from Quechua to Spanish as needed, and transcribed into Microsoft Word. Analysis was done by a trained medical anthropologist with Atlas.ti software.\(^{56}\)

**RESULTS**

**Results of Intervention Monitoring**

Attendance by CHWs and CSs was 82% or better for 5 of 6 workshop topics in the monthly trainings offered at the 22 PHC facilities. Additional small-group training sessions were run by CSs for their respective CHWs once or twice a month within communities for reinforcement of learning. CHWs covered missed workshops during these small group sessions or through a CS visiting them at home to provide personalized training.

Verbally applied knowledge tests were given to CHWs before and after completing each training module topic. Pretest scores of CHWs averaged 40% and improved substantially to about 80% on posttests. Experimentally trained CHW had posttest scores much higher than controls on the growth and nutrition and the diarrhea modules, which were 90% and 96% for the experimental group, respectively; the CHW control group scored 82% and 74%, respectively. CHW workshop attendance and pre-post test scores are reported in the Supplement.

**Baseline and Follow-Up Household Survey Results**

**Comparability of Study Groups on Demographic Characteristics**

Mothers were comparable between study groups by age, parity, and education. More control group mothers worked for cash or barter in the follow-up survey compared with experimental group
mothers. Children were comparable between study groups in both surveys by age, sex, mean birthweight, and proportion with low birth weight (Table 3).

**Reported Household Visits by Community Health Workers**

As reported in the endline household survey, the proportion of mothers who received 1 or more home visits from a CHW during pregnancy or after birth was 63.1% in experimental clusters and 60.5% in control clusters, with a similar distribution of number of visits by study group. Mothers in the experimental group were more likely to receive 1 or more CHW visits compared with those in the control group among those with any primary school education (72.0% experimental versus 66.4% control, \( P < .01 \)). Among those visited,

### TABLE 3. Demographic Characteristics of Mothers and Children, by Study Group and Survey, Cluster-Randomized Controlled Trial on Sharing Histories CHW Training Methodology, Huánuco Region, Peru

<table>
<thead>
<tr>
<th>Demographic Characteristics</th>
<th>Baseline Survey 2010</th>
<th>Study Group</th>
<th>Final Survey 2014</th>
<th>Study Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Experimental</td>
<td>Control</td>
<td>Significance</td>
<td>Experimental</td>
</tr>
<tr>
<td></td>
<td>(n=308)</td>
<td>(N=298)</td>
<td></td>
<td>(n=290)</td>
</tr>
<tr>
<td>Mothers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, years, mean (SD)</td>
<td>26.9 (7.8)</td>
<td>27.2 (9.2)</td>
<td>.62</td>
<td>27.1 (7.8)</td>
</tr>
<tr>
<td>Number of children, mean (SD)</td>
<td>2.6 (1.9)</td>
<td>2.8 (2.2)</td>
<td>.18</td>
<td>2.7 (1.8)</td>
</tr>
<tr>
<td>Distribution of number of children, %</td>
<td>.14</td>
<td>.07</td>
<td>.35</td>
<td>.60</td>
</tr>
<tr>
<td>1</td>
<td>37.5</td>
<td>33.0</td>
<td></td>
<td>35.9</td>
</tr>
<tr>
<td>2</td>
<td>22.6</td>
<td>27.9</td>
<td></td>
<td>22.8</td>
</tr>
<tr>
<td>3–4</td>
<td>26.2</td>
<td>21.4</td>
<td></td>
<td>22.8</td>
</tr>
<tr>
<td>5–12</td>
<td>13.6</td>
<td>17.7</td>
<td></td>
<td>18.6</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
<td>100.0</td>
</tr>
<tr>
<td>Education, years, mean (SD)</td>
<td>4.4 (3.7)</td>
<td>4.7 (3.6)</td>
<td>.35</td>
<td>5.4 (3.9)</td>
</tr>
<tr>
<td>Distribution of maternal educational level, % (n)</td>
<td>.60</td>
<td>.37</td>
<td>.35</td>
<td>.60</td>
</tr>
<tr>
<td>No education or cannot read (illiterate)</td>
<td>31.5 (97)</td>
<td>32.3 (96)</td>
<td></td>
<td>24.8 (72)</td>
</tr>
<tr>
<td>Any primary education (literate)</td>
<td>46.2 (142)</td>
<td>42.5 (127)</td>
<td></td>
<td>40.7 (118)</td>
</tr>
<tr>
<td>Any secondary or more (literate)</td>
<td>22.3 (69)</td>
<td>25.2 (75)</td>
<td></td>
<td>34.5 (100)</td>
</tr>
<tr>
<td>Total</td>
<td>100 (308)</td>
<td>100 (298)</td>
<td></td>
<td>100 (290)</td>
</tr>
<tr>
<td>Works for cash or barter, %</td>
<td>12.8</td>
<td>17.3</td>
<td>.25</td>
<td>34.1</td>
</tr>
<tr>
<td>Children</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age in months, mean (SD)</td>
<td>11.4 (6.7)</td>
<td>10.9 (6.8)</td>
<td>.39</td>
<td>11.0 (7.0)</td>
</tr>
<tr>
<td>Sex of child, female, %</td>
<td>47.7</td>
<td>47.3</td>
<td>.49</td>
<td>50.7</td>
</tr>
<tr>
<td>Birth weight (g), mean (SD)</td>
<td>3,045 (471)</td>
<td>3,042 (285)</td>
<td>.92</td>
<td>3,051 (444)</td>
</tr>
<tr>
<td>For illiterate mothers</td>
<td>3,063 (454)</td>
<td>2,953 (494)</td>
<td>.13</td>
<td>2,980 (448)</td>
</tr>
<tr>
<td>For literate mothers</td>
<td>3,042 (479)</td>
<td>3,080 (436)</td>
<td>.41</td>
<td>3,074 (442)</td>
</tr>
<tr>
<td>Low birth weight (&lt;2,500 g)</td>
<td>10.2</td>
<td>7.2</td>
<td>.30</td>
<td>9.8</td>
</tr>
<tr>
<td>For illiterate mothers</td>
<td>10.3</td>
<td>10.2</td>
<td>.98</td>
<td>9.9</td>
</tr>
<tr>
<td>For literate mothers</td>
<td>10.2</td>
<td>6.7</td>
<td>.20</td>
<td>9.8</td>
</tr>
</tbody>
</table>

*Abbreviation: CHW, community health worker; SD, standard deviation.*
mothers in both study groups received an average of 5.3 visits from a CHW before and/or after their most recent pregnancy, with a range of 1 to 27 visits. (Table 4).

**Changes in Knowledge and Practices of Study Mothers**

**Knowledge of Key Danger Signs.** Maternal spontaneous knowledge of 2 or more danger signs on each of 4 topics (pregnancy, birth, postpartum, and newborns) increased significantly in both study groups from baseline to endline. Improvements from baseline to endline were greater in the experimental group than in the control group regarding danger signs in pregnancy and postpartum. The control group of mothers had greater improvement in knowledge of danger signs in newborns (Table 5).

**Micronutrient Consumption.** Study mothers increased their consumption of iron tablets for the standard minimum of 3 or more months during pregnancy at an approximate rate of 50% increase in both study groups. Study children consumed micronutrient powder (Sprinkles) added to food in the previous 24 hours in approximately 60% of both study groups at endline. No micronutrients were taken at baseline by either group (Table 5).

**Key Breastfeeding Practices.** Nearly 90% of all study children were breastfeeding at the time of both surveys. At baseline, three-fourths of study cases had initiated breastfeeding within 1 hour of birth, but this declined to two-thirds in both groups at endline. Exclusive breastfeeding in children 0–5 months of age increased significantly from baseline to endline by 14.9 points in the experimental group.

### Table 4. Home Visits From CHWs Received by Mothers at Endline Survey, 2014, by Study Group, Huánuco Region, Peru

<table>
<thead>
<tr>
<th>Receipt by Mothers of CHW Home Visits</th>
<th>Experimental (n=290)</th>
<th>Control (n=263)</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mothers who received 1 or more CHW home visits, %</td>
<td>63.1</td>
<td>60.5</td>
<td>.29</td>
</tr>
<tr>
<td>Distribution of mothers who received 1 or more home visits by educational level, %</td>
<td>23.5</td>
<td>18.9</td>
<td>.59</td>
</tr>
<tr>
<td>No education or cannot read (illiterate)</td>
<td>23.5</td>
<td>18.9</td>
<td>.59</td>
</tr>
<tr>
<td>Any primary education and can read (literate)</td>
<td>46.4</td>
<td>49.7</td>
<td>.21</td>
</tr>
<tr>
<td>Any secondary or more education (literate)</td>
<td>30.1</td>
<td>31.4</td>
<td>.55</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>.55</td>
</tr>
<tr>
<td>Mothers who received 1 or more CHW visits received within each educational level, %</td>
<td>59.7</td>
<td>56.6</td>
<td>.43</td>
</tr>
<tr>
<td>No education or cannot read (illiterate)</td>
<td>59.7</td>
<td>56.6</td>
<td>.43</td>
</tr>
<tr>
<td>Any primary education and can read (literate)</td>
<td>72.0</td>
<td>66.4</td>
<td>.21</td>
</tr>
<tr>
<td>Any secondary or more education (literate)</td>
<td>55.0</td>
<td>54.9</td>
<td>.55</td>
</tr>
<tr>
<td>Number of CHW home visits received (range 1–27), mean (SD)</td>
<td>5.28 (4.6) (N=180)</td>
<td>5.27 (4.3) (N=159)</td>
<td>.99</td>
</tr>
<tr>
<td>Number of CHW home visits received within each educational level (range 0–27), mean (SD)</td>
<td>3.28 (4.5)</td>
<td>3.19 (4.2)</td>
<td>.81</td>
</tr>
<tr>
<td>No education or cannot read (illiterate) (range 0–27)</td>
<td>2.65 (3.9)</td>
<td>2.98 (4.1)</td>
<td>.65</td>
</tr>
<tr>
<td>Any primary education and can read (literate) (range 0–27)</td>
<td>3.51 (4.0)</td>
<td>3.46 (4.3)</td>
<td>.93</td>
</tr>
<tr>
<td>Any secondary or more education (literate) (range 0–27)</td>
<td>3.45 (5.3)</td>
<td>2.94 (4.2)</td>
<td>.47</td>
</tr>
<tr>
<td>Distribution of number of CHW visits received by mothers, %</td>
<td>37.9</td>
<td>39.2</td>
<td>.88</td>
</tr>
<tr>
<td>0</td>
<td>37.9</td>
<td>39.2</td>
<td>.88</td>
</tr>
<tr>
<td>1 or 2</td>
<td>19.3</td>
<td>19.0</td>
<td>.88</td>
</tr>
<tr>
<td>3–5</td>
<td>22.4</td>
<td>20.2</td>
<td>.88</td>
</tr>
<tr>
<td>6–10</td>
<td>13.1</td>
<td>15.2</td>
<td>.88</td>
</tr>
<tr>
<td>11–27</td>
<td>7.2</td>
<td>6.1</td>
<td>.88</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>.88</td>
</tr>
</tbody>
</table>

Abbreviations: CHW, community health worker, SD, standard deviation.
| TABLE 5. Changes in Maternal Knowledge and Practice, by Study Group, Huánuco Region, Peru |
|---------------------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                                                | Baseline Survey 2010 |                  | Endline Survey 2014 |                  | Baseline to Endline Differences |
|                                                | Experimental N=308 | Control N=298 | Experimental N=290 | Control N=263 |                                  |
| Proportion of mothers with spontaneous correct report of at least 2 danger signs, % | | | | | | | | | |
| Danger signs in pregnancy                      | 39.3 | 42.6 | .28 | 73.8 | 74.5 | .46 | +34.5 | +27.9 |
| Danger signs during birth                       | 11.0 | 7.7 | .10 | 27.6 | 24.7 | .25 | +16.6 | +17.0 |
| Danger signs in postpartum                     | 20.5 | 23.2 | .24 | 38.3 | 35.7 | .30 | +17.8 | +12.5 |
| Danger signs in newborns                       | 23.1 | 18.1 | .08 | 61.0 | 61.6 | .48 | +37.9 | +43.5 |
| Micronutrient consumption, %                   | | | | | | | | | |
| Mothers consuming iron tabs 3+ months last pregnancy | 51.0 | 46.6 | .16 | 77.2 | 74.1 | .41 | +26.2 | +27.5 |
| Children with micronutrients added to food in past 24 hours | 0.0  | 0.3 | .49 | 57.9 | 61.2 | .24 | +57.9 | +60.9 |
| Children with vitamin A supplement in past 6 months | 47.1 | 46.3 | .77 | 21.3 | 19.2 | .28 | -25.8 | -27.1 |
| Proportion of children with nutritional pattern, % | | | | | | | | | |
| Currently breastfeeding                         | 89.6 | 87.2 | .36 | 88.3 | 90.1 | .29 | -1.3 | +2.9 |
| Early breastfeeding within 1 hour of birth       | 77.6 | 72.8 | .11 | 69.0 | 66.5 | .30 | -8.6 | -6.3 |
| Exclusive breastfeeding, 0-5 months             | 71.8 | 83.5 | .06 | 86.6 | 90.2 | .35 | +14.8 | +6.7 |
| Food consumption in past 24 hours for children aged 6-23 months, % | | | | | | | | | |
| Iron-rich foods                                 | 53.2 | 53.9 | .48 | 93.8 | 93.1 | .47 | +40.6 | +39.2 |
| Animal protein                                  | 32.5 | 37.4 | .16 | 47.1 | 49.0 | .39 | +14.6 | +11.6 |
| Minimum meal frequency                          | 69.6 | 68.0 | .40 | 94.2 | 89.6 | .06 | +24.6 | +21.6 |
| Minimum food diversity                          | 56.1 | 58.4 | .34 | 79.3 | 75.7 | .23 | +19.6 | +17.3 |
| Household water, sanitation, and hygiene practices by maternal literacy, % | | | | | | | | | |
| No animals (except pets) live inside house      | Illiterate 45.0 | 50.0 | .31 | 52.8 | 64.2 | .14 | +7.8 | +14.2 |
| Uses correct treatment for drinking water       | Literate 50.6 | 56.9 | .14 | 65.6 | 58.6 | .08 | +15.0 | +1.7 |
| Mother washes hands after defecating           | Illiterate 80.2 | 89.5 | .17 | 97.2 | 79.2 | .00 | +17.0 | -10.3 |
| Soap is available for hand washinga             | Literate 86.1 | 85.9 | .29 | 92.2 | 96.2 | .06 | +6.1 | +10.3 |
| Uses safe water source                          | Illiterate 28.1 | 29.5 | .48 | 23.6 | 34.0 | .14 | -4.5 | +4.5 |
| Improved cook stove installed in past 4 years   | Literate 50.7 | 43.2 | .08 | 50.5 | 39.0 | .01 | -0.2 | -4.2 |
| Does not use wood or dried dung as cook fuel    | Illiterate 49.0 | 43.2 | .26 | 66.7 | 63.5 | .43 | +17.7 | +20.3 |
| Receipt of government health and social services, % | | | | | | | | | |
| Infant food supplementation program            | 93.8 | 92.3 | .28 | 0.7 | 1.1 | .45 | -93.1 | -91.2 |
| Conditional cash transfer program (Juntos)       | 52.6 | 53.0 | .49 | 53.8 | 52.9 | .45 | +1.2 | -0.1 |

Continued
TABLE 5. Continued

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Baseline Survey 2010</th>
<th>Endline Survey 2014</th>
<th>Baseline to Endline Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Experimental N=308</td>
<td>Control N=298</td>
<td>Significance</td>
</tr>
<tr>
<td>Municipal Glass of Milk program</td>
<td>76.0</td>
<td>76.2</td>
<td>0.52</td>
</tr>
<tr>
<td>Child antiparasite treatment in past 6 months</td>
<td>12.1</td>
<td>15.4</td>
<td>0.13</td>
</tr>
<tr>
<td>Participation of mothers in women’s groups with discussion of child health and nutrition*</td>
<td>46.4</td>
<td>44.3</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Not assessed at baseline.

(71.8% to 86.7%) ($\chi^2(1,N=602)= 20.4; P<.01$), but by only 6.7 points in the control group (83.5% to 90.2%) ($\chi^2(1,N=598)=5.6; P<.05$) (Table 5).

Key Child Feeding Practices. Children’s consumption of iron-rich foods notably increased in both study groups from 53% at baseline to 93% at endline. Animal protein consumption by children also increased significantly in both study groups but reached less than 50% at endline (Table 5). Minimum meal frequency, established as 3 meals per day on average according to age and breastfeeding status,54 was 2.2 meals per day for infants 6–11 months at baseline increasing to 3.5 per day at endline. The experimental group had a significantly higher percentage of minimum meal frequency than controls at endline (94.2% experimental versus 89.6% control) ($\chi^2(1,N=584)=4.1; P<.05$) (Table 5). Minimum dietary diversity by breastfeeding status54 showed significant increases in both study groups but differences between study groups were nonsignificant at endline (Table 5).

Key Water, Sanitation, and Hygiene Practices. WASH practices at baseline were generally much better among literate mothers than nonliterate mothers in both study groups. At endline, several key WASH practices differed significantly within strata of maternal literacy. Among literate mothers, no animals living inside the home (except pets) and mother washing hands after defecation were more frequent in the experimental group compared to the control group. Among illiterate mothers, correct treatment of drinking water and installation of an improved cook stove in the past 4 years were significantly more frequent in the experimental versus control group (Table 5).

Government Health and Social Services Received by Mothers and Children. Study groups were comparable on receipt of government services for mothers and children that might affect child growth. Receipt of a government food supplement (instant fortified weaning food) by children aged 6–23 months was nearly universal at baseline but null at endline because the food program was discontinued in 2012. One-half of all study mothers received $30 per month as a conditional cash transfer (Juntos Program). The municipal Glass of Milk program provided a daily milk ration to about two-thirds of all mothers with a child 0–5 months of age and to children aged 6–23 months. Government distribution of antiparasitic medication and vitamin A supplements declined during the project period. One-half of all mothers participated in a women’s group in which health and nutrition topics were discussed (Table 5).

Baseline to Endline Changes in Child Stunting

Changes in Stunting by Demographic Characteristics

The baseline prevalence of child stunting was 34%–35% in both study groups, unadjusted for clustering. Stunting was reduced in experimental clusters by 4.1% from baseline (34.4%) to endline (30.3%), while stunting in control clusters plateaued from baseline (35.3%) to endline (35.0%) (Table 6). The difference at endline is not significant.

At both baseline and endline, stunting was much lower in children of literate mothers compared with children of illiterate mothers, in children aged 0–11 months compared with those aged 12–23 months, in girls compared with boys, and in those with normal birthweight (2,500 g or more) compared with those with low birth weight (≤2,500 g) (Table 6).

At endline, stunting prevalence among children of literate mothers with any primary or secondary education was significantly lower in the experimental group at 24.8% versus 33.8% among controls. Among children with normal birth weight >2,500 g, stunting was lower in the experimental...
In accordance with the findings of significantly greater reduction in stunting in children of literate mothers in the experimental group, we found that these mothers, compared with control group peers, had reported the following more frequently: no animals kept within the home (65.5% experimental versus 58.6% control, \( P < .05 \)); handwashing after defecation (50.5% experimental versus 39.0% control, \( P < .05 \)); and provision of a minimum number of feeds per day to their child (94.2% experimental versus 89.6% control, \( P < .05 \)) (Table 5).

### Changes in Stunting in the per Protocol Subgroup

 Mothers who received 1 or more CHW visits were significantly less educated than mothers who reported no visits. Nevertheless, the prevalence of child stunting did not differ between the visited and non-visited groups (Table 7).

In the per protocol analysis of mothers who had been visited by CHWs, stunting was present in 29.0% of children of mothers served by an experimentally trained CHW compared with 38.4% in children of mothers served by control CHWs (\( P = .04 \)). Among literate mothers, these percentages were 22.9% and 36.4%, respectively, for experimental and control groups (\( P = .01 \)) (Table 7 and Figure 3). Literate mothers were 78% of the study population at endline.

### Predictors of Child Stunting

Bivariate regressions of potential predictors of child stunting, including type of CHW learning method (SH versus control) and potential covariates, are shown in Table 8, adjusted for clustering. The experimental CHW learning method and maternal literacy were found to interact in relation to stunting. That is, learning method was associated with a decrease in stunting among children of literate mothers, but not among children of illiterate mothers. The interaction term had a significant independent association with child stunting (Beta=0.75; 95% confidence interval=0.20, 1.30; \( P < .00 \)) (Table 8). Covariates that had bivariate association with stunting significant at \( P < .20 \) included child age in months, sex of child, low birth weight (<2,500 g), birth weight in grams, child consumed the minimum diversity of food in the previous 24 hours, number of months that mother took iron tablets during pregnancy, child consumed multi-micronutrient + iron supplement in the past 24 hours, household has an improved cookstove installed in the previous 4 years, and child had parasite treatment in the past 6 months. Variables that were colinear

The experimental CHW learning method and maternal literacy were found to interact in relation to stunting.

<table>
<thead>
<tr>
<th>TABLE 6. Changes in Prevalence of Growth Stunting in Children Aged 0–23 Months by Demographic Characteristics, Study Group, and Survey, Huánuco Region, Peru</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline Survey 2010</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>All study children</td>
</tr>
<tr>
<td>Mother’s educational level</td>
</tr>
<tr>
<td>No education or cannot read</td>
</tr>
<tr>
<td>Any primary or secondary</td>
</tr>
<tr>
<td>Child’s age</td>
</tr>
<tr>
<td>0–11 months</td>
</tr>
<tr>
<td>12–23 months</td>
</tr>
<tr>
<td>Child’s sex</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Child birth weight</td>
</tr>
<tr>
<td>( \leq 2,500 \text{ g} )</td>
</tr>
<tr>
<td>( &gt; 2,500 \text{ g} )</td>
</tr>
</tbody>
</table>
### TABLE 7. Changes in Prevalence of Growth Stunting in Children Aged 0–23 Months by Receipt of 1 or More CHW Visits, by Maternal Literacy and Study Group at Endline, Huánuco Region, Peru

<table>
<thead>
<tr>
<th>All mothers</th>
<th>With CHW Visit</th>
<th>Without CHW Visit</th>
<th>Significance</th>
<th>With CHW Visit</th>
<th>Control</th>
<th>Significance</th>
<th>Without CHW Visit</th>
<th>Control</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child stunting, %</td>
<td>33.3</td>
<td>31.3</td>
<td>.34</td>
<td>29.0</td>
<td>38.4</td>
<td>.04</td>
<td>37.2</td>
<td>29.8</td>
<td>.38</td>
</tr>
<tr>
<td>Maternal education in years, mean (SD)</td>
<td>5.3 (3.5)</td>
<td>6.0 (3.9)</td>
<td>.02</td>
<td>5.1 (3.6)</td>
<td>5.6 (3.4)</td>
<td>.20</td>
<td>5.9 (4.2)</td>
<td>6.1 (3.7)</td>
<td>.71</td>
</tr>
<tr>
<td>N</td>
<td>342</td>
<td>211</td>
<td></td>
<td>183</td>
<td>159</td>
<td>107</td>
<td>104</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illiterate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child stunting, %</td>
<td>47.9</td>
<td>38.5</td>
<td>.19</td>
<td>48.8</td>
<td>46.7</td>
<td>.52</td>
<td>44.8</td>
<td>30.4</td>
<td>.22</td>
</tr>
<tr>
<td>Maternal education in years, mean (SD)</td>
<td>0.7 (1.3)</td>
<td>1.2 (1.4)</td>
<td>.08</td>
<td>0.7 (1.3)</td>
<td>0.8 (1.2)</td>
<td>.67</td>
<td>1.0 (1.4)</td>
<td>1.4 (1.4)</td>
<td>.38</td>
</tr>
<tr>
<td>N</td>
<td>73</td>
<td>52</td>
<td></td>
<td>43</td>
<td>30</td>
<td>29</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Literate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child stunting, %</td>
<td>29.4</td>
<td>28.9</td>
<td>.51</td>
<td>22.9</td>
<td>36.4</td>
<td>.01</td>
<td>28.2</td>
<td>29.6</td>
<td>.49</td>
</tr>
<tr>
<td>Maternal education in years, mean (SD)</td>
<td>6.5 (2.9)</td>
<td>7.6 (3.1)</td>
<td>.00</td>
<td>6.4 (3.0)</td>
<td>6.7 (2.7)</td>
<td>.49</td>
<td>7.8 (3.3)</td>
<td>7.5 (2.9)</td>
<td>.58</td>
</tr>
<tr>
<td>N</td>
<td>269</td>
<td>159</td>
<td></td>
<td>140</td>
<td>129</td>
<td>78</td>
<td>81</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CHW, community health worker, SD, standard deviation.

### FIGURE 3. Changes in Stunting in Children Aged 0–23 Months Before and After a Community Health Worker Training Intervention Comparing Experimental and Control Groups by Maternal Literacy, Huánuco Region, Peru
<table>
<thead>
<tr>
<th>Predictors of Child Stunting&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Estimate Beta</th>
<th>95% Confidence Interval</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental group</td>
<td>0.21</td>
<td>−0.31, 0.74</td>
<td>.43</td>
</tr>
<tr>
<td>Mother is literate</td>
<td>0.64</td>
<td>0.30, 0.98</td>
<td>&lt;.00</td>
</tr>
<tr>
<td>Interaction: experimental group × maternal literacy</td>
<td>0.75</td>
<td>0.20, 1.30</td>
<td>&lt;.00</td>
</tr>
<tr>
<td>Maternal characteristics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of children born in mother’s lifetime (1–12)</td>
<td>−0.12</td>
<td>−0.24, 0.01</td>
<td>.07</td>
</tr>
<tr>
<td>Mother has remunerated work (0, 1)</td>
<td>0.00</td>
<td>−0.34, 0.34</td>
<td>.99</td>
</tr>
<tr>
<td>Child characteristics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (0–23 months)</td>
<td>−0.08</td>
<td>−0.10, −0.05</td>
<td>&lt;.00</td>
</tr>
<tr>
<td>Female child (0, 1)</td>
<td>0.63</td>
<td>0.26, 0.99</td>
<td>&lt;.00</td>
</tr>
<tr>
<td>Low birth weight (&lt;2,500 g)</td>
<td>−0.75</td>
<td>1.32, −0.72</td>
<td>.01</td>
</tr>
<tr>
<td>Birth weight (g)</td>
<td>.00</td>
<td>.00, .00</td>
<td>&lt;.00</td>
</tr>
<tr>
<td>Child feeding variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Receives breastmilk within 1 hour of birth (0, 1)</td>
<td>−0.19</td>
<td>−0.62, 0.24</td>
<td>.39</td>
</tr>
<tr>
<td>Meets minimum food diversity (0, 1)</td>
<td>−0.57</td>
<td>−1.18, 0.054</td>
<td>.07</td>
</tr>
<tr>
<td>Meets minimum meal frequency (0, 1)</td>
<td>−0.17</td>
<td>−0.86, 0.53</td>
<td>.64</td>
</tr>
<tr>
<td>Consumes iron-rich foods in past 24 hours (0, 1)</td>
<td>−0.35</td>
<td>−1.47, 0.78</td>
<td>.55</td>
</tr>
<tr>
<td>Consumes animal food source past 24 hours (0, 1)</td>
<td>0.02</td>
<td>−0.55, 0.60</td>
<td>.94</td>
</tr>
<tr>
<td>Micronutrient consumption</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother took iron during last pregnancy (0–9 months)</td>
<td>0.09</td>
<td>−0.02, 0.17</td>
<td>.06</td>
</tr>
<tr>
<td>Child consumed Sprinkles® in past 24 hours (0, 1)</td>
<td>−0.45</td>
<td>−0.86, −0.04</td>
<td>.03</td>
</tr>
<tr>
<td>Vitamin A capsule taken by child in past 6 months (0, 1)</td>
<td>−0.04</td>
<td>−0.56, 0.49</td>
<td>.89</td>
</tr>
<tr>
<td>Water, sanitation, and hygiene practices</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother washes hands after defecation (0, 1)</td>
<td>0.10</td>
<td>−0.17, 0.36</td>
<td>.46</td>
</tr>
<tr>
<td>Soap, ash, or detergent used to wash hands (0, 1)</td>
<td>0.03</td>
<td>−0.19, 0.24</td>
<td>.80</td>
</tr>
<tr>
<td>Household has safe water source (0, 1)</td>
<td>−0.35</td>
<td>−0.85, 0.15</td>
<td>.17</td>
</tr>
<tr>
<td>Drinking water is treated correctly (0, 1)</td>
<td>0.20</td>
<td>−0.33, 0.73</td>
<td>.45</td>
</tr>
<tr>
<td>Improved cook stove installed in past 4 years (0, 1)</td>
<td>−0.44</td>
<td>−0.85, −0.03</td>
<td>.03</td>
</tr>
<tr>
<td>Non-pet animals do not live in the home (0, 1)</td>
<td>0.14</td>
<td>−0.26, 0.54</td>
<td>.50</td>
</tr>
<tr>
<td>Government health and social services</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parasite treatment for child in past 6 months (0, 1)</td>
<td>−0.90</td>
<td>−1.43, −0.37</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Mother in participatory women’s group (0, 1)</td>
<td>−0.08</td>
<td>−0.47, 0.31</td>
<td>.68</td>
</tr>
<tr>
<td>Glass of Milk daily ration for child (0, 1)</td>
<td>−0.08</td>
<td>−0.48, 0.33</td>
<td>.72</td>
</tr>
<tr>
<td>Junto's cash transfer received by mother (0, 1)</td>
<td>−0.10</td>
<td>−0.46, 0.27</td>
<td>.59</td>
</tr>
</tbody>
</table>

Abbreviation: CHW, community health worker.

<sup>a</sup> Adjusted for clustering.

<sup>b</sup> Outcome variable: stunted=1, not stunted=0.

<sup>c</sup> Multimicronutrients with iron.
with the outcome (birth weight in grams, child age in months) were not included in the multivariate model.

**Effect of the Experimental Training Intervention on Stunting**

**DID Analysis**

We conducted a DID analysis for all children, and for stratified subgroups of maternal literacy. Results in Table 9 show that for all children, the mean HAZ change was not significantly different between experimental and control clusters ($P = .469$). However, in the subgroup of literate mothers, the mean HAZ improved by an average of 1.03 points on the Z-scale in experimental clusters as compared to control clusters between baseline and endline, with a significance level of $P = .059$. For the subgroup of illiterate mothers, the changes between baseline and endline for experimental versus control clusters were not significantly different.

We then used DID to test whether the intervention’s effect differed by maternal literacy as an interaction effect. We found that the difference of the mean HAZ from baseline to endline in each of the matched clusters between children of high- and low-literacy mothers in the experimental versus the control group was significant at $P = .003$, using a paired t-test. This finding demonstrated a significant interaction effect of the experimental training method on HAZ by level of maternal literacy.

**GEE Analysis**

Table 10 shows the main variables submitted to the multivariate model using GEE with adjustment for clustering: inclusion in the experimental group, maternal literacy, and the interaction term (experimental group × high maternal literacy) plus covariates. Multivariate results showed that the interaction term had a significant association with stunting ($B = .77; 95\% CI = 0.23, 1.31; P < .00$). In other words, SH was significantly associated with reduced stunting among literate mothers but not among illiterate mothers. This effect was adjusted by significant covariates associated with stunting: child parasite treatment in the past 6 months, child consumed multi-micronutrients with iron in the past 24 hours, and improved cook stove installed in the past 4 years. Safe water source had a borderline association with stunting.

**Results of Close-Ended Interviews With CHWs and CSs**

Results of interviews with 50 randomly selected CHWs from each study group (n=100) and all 27 CSs at project end showed that experimentally trained CHWs and CSs felt more positive about their training and learning than those trained with the control method.

Experimental CHWs felt more capable of identifying danger signs in mothers and children during the first 1,000 days than control CHWs (38\% experimental versus 20\% control, $P < .05$).

When CHWs were asked about activities they liked best or found most interesting in the workshops, experimental CHWs liked all the training activities more frequently than control CHWs. Training activities liked best were the use of the SH method (86\% experimental versus 60\% control, $P < .01$), the pretest (74\% experimental versus 52\% control, $P < .05$), and the posttest (78\% experimental versus...

---

**Table 9.** DID Analysis of Mean HAZ in 533 Children Aged 0–23 Months, Baseline to Endline, for All Children and for Subgroups of Children Stratified by Maternal Literacy, Huánuco Region, Peru

<table>
<thead>
<tr>
<th>Group</th>
<th>Control</th>
<th>Experimental</th>
<th>Paired t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean HAZ, Baseline, Mean (SD)</td>
<td>Mean HAZ, Endline, Mean (SD)</td>
<td>Diff. of Means, Mean (SD)</td>
</tr>
<tr>
<td>All children</td>
<td>-1.55 (.46)</td>
<td>-1.55 (.66)</td>
<td>-0.0002 (.36)</td>
</tr>
<tr>
<td>Stratified by literacy of mother</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illiterate</td>
<td>-1.79 (.51)</td>
<td>-1.57 (.93)</td>
<td>0.22 (.59)</td>
</tr>
<tr>
<td>Literate</td>
<td>-0.80 (1.13)</td>
<td>-1.53 (0.59)</td>
<td>-0.74 (1.01)</td>
</tr>
</tbody>
</table>

Abbreviations: df, degrees of freedom; DID, difference-in-differences; HAZ, height-for-age-Z-scores; SD, standard deviation.
Community supervisor meeting with small group of community health workers to reinforce training received from health personnel trainer that month.
Credit: ©2013 Elie Gardner

### TABLE 10. Generalized Estimating Equationsa Multivariate Model for Predictors of Child Stunting With 553 Children Aged 0–23 Months, Huánuco Region, Peru

<table>
<thead>
<tr>
<th>Predictors of Child Stuntingb</th>
<th>Estimates</th>
<th>95% Wald Confidence Interval</th>
<th>Hypothesis Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta</td>
<td>SE</td>
<td>Lower</td>
</tr>
<tr>
<td>Community health worker training intervention</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental group (0, 1)</td>
<td>−0.27</td>
<td>0.36</td>
<td>0.97</td>
</tr>
<tr>
<td>Mother is literate (0, 1)</td>
<td>0.22</td>
<td>0.11</td>
<td>0.00</td>
</tr>
<tr>
<td>Interaction: experimental group × maternal literacy</td>
<td>0.77</td>
<td>0.27</td>
<td>0.23</td>
</tr>
<tr>
<td>Child nutrition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child consumed Sprinklesc past 24 hours (0, 1)</td>
<td>−0.41</td>
<td>0.21</td>
<td>−0.81</td>
</tr>
<tr>
<td>Water, sanitation, and hygiene</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Safe water source (0, 1)</td>
<td>−0.43</td>
<td>0.25</td>
<td>−0.92</td>
</tr>
<tr>
<td>Improved cook stove installed past 4 years (0, 1)</td>
<td>−0.49</td>
<td>0.21</td>
<td>−0.91</td>
</tr>
<tr>
<td>Government health and social services</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parasite treatment for child in past 6 months (0, 1)</td>
<td>−.85</td>
<td>.34</td>
<td>−1.52</td>
</tr>
<tr>
<td>Intercept</td>
<td>1.22</td>
<td>0.35</td>
<td>0.54</td>
</tr>
</tbody>
</table>

Goodness of fitd with corrected quasi-likelihood under independence model criterion (QICC) 672.97

Abbreviations: df, degrees of freedom; SE, standard error.

a Adjusted for clustering.
b Outcome variable: stunted=1, not stunted=0.
c Multimicronutrients with iron.
d Information criteria are in smaller-is-better form.
48% control, P < .01). Compared with control CHWs, experimental CHWs more frequently liked or found interesting the explanation of the flip charts (92% experimental versus 82% control) and use of participatory dynamics for evaluation of learning at the end of each learning session (86% experimental versus 72% control).

When asked what they had found least entertaining or interesting in the training workshops, control CHWs were more likely to mention disliking all the activities than the experimental CHWs: Sharing Histories (0% experimental versus 8% control, P < .05); pretest (14% experimental versus 38% control, P < .01); posttest (6% experimental versus 16% control, P = .11); explanation of the flip chart (2% experimental versus 12% control, P < .05); and participatory dynamics for evaluation (12% experimental versus 26% control, P = .07).

Among 27 CSs interviewed, decentralized group meetings held by experimentally trained CSs with their assigned CHWs were more likely to last over 2 hours, compared with meetings held by control CSs (64% experimental versus 31% control, P < .05)

In addition, more experimentally trained CSs reported that community leaders supported the work of CHWs in communities than did control CSs, in the following ways: leaders reached out to the municipality to implement actions for women and children (79% experimental versus 69% control); leaders prepared community work plans with activities to protect women and children (71% experimental versus 58% control); leaders gave orders or made community resolutions to encourage families to adopt health practices (36% experimental versus 7% control); and leaders helped women to seek care in primary care facilities or hospitals (29% experimental versus 15% control).

Results of Qualitative In-Depth Interviews

Postintervention qualitative interviews of CHWs, CSs, and trainers regarding use of SH suggested mechanisms through which the method could have promoted health behavior change first in female CHWs and then in mothers taught by trained CHWs.

To begin with, SH seemed to echo traditional conceptions in this population that knowledge is based on experience.

CHWs noted that after listening to each other’s histories during a training session, seeing a belief or practice listed on paper marked in red for “possibly dangerous effect on health” motivated them to learn why it could be dangerous. After discussing the traditional practice, they discussed what the healthy behavior should be instead, with picture cards (flip chart) later serving to reinforce what the new behaviors should be. (Names of persons cited have been changed to protect privacy.)

_They had us remember how we used to do things and then they taught us._ —CHW Jenny

CHWs’ memories of past experiences were frequently similar because of shared cultural practices. Certain practices that could be harmful, such as home birthing, were often collective practice. Once the practice was mentioned, the group could discuss it openly without recrimination, with a shared acceptance of new learning by CHWs. Hearing about their peers’ differences in practice also provided CHWs with motivation to learn.

_One [CHW] knows...and takes the mother to the hospital, and others [CHWs] don’t know. I felt a greater urge to want to learn._ —CHW Maria

Some memories were emotionally charged, such as pregnancy or birth complications or the death of a child. When strong emotions and tears were expressed, empathetic responses from fellow learners worked to strengthen interpersonal relationships.

_...Some were born badly, sick... I thought, I thank God because my daughter was born well. On the other hand, there are mothers who suffered. I thought, how is that possible. Sometimes you don’t know._ —CHW Diana

By sharing histories, CHWs seemed to become self-motivated to avoid repetition of their own prior erroneous behavior. CHWs said they felt more confident in their ability to explain new practices to other mothers through the lens of local belief systems; they were able to share newfound knowledge with greater force of conviction.

_It is a good lesson. Because sharing histories you can find out...how that person is living. And you can help them._ ..._You give them confidence and they tell you and a solution can be found. Because by not telling your problems, you don’t find out anything about anybody. But if they tell you their histories, their problems, yes. And you can help them._ —CHW Jenny

CHWs spontaneously shared their own experiences with the mothers under their care, and in turn asked them to share their experiences, thereby promoting empathic connections to strengthen the social relationship, bonding, and trust needed for influencing traditional knowledge.
I feel better because now I know...what it is to teach mothers. Before I didn’t know. So, on that side I feel good and I have learned to teach. —CHW Diana

Health personnel who learned to train CHWs with the SH method initially struggled with the new way of training, but soon learned to appreciate the method.

At first it was difficult but soon it seemed natural. By sharing histories, they take interest in the topic and the new knowledge sticks with them. For us it is easier to teach them like that. —Trainer María

Some examples of cultural beliefs that were expressed by CHWs during the history-sharing sessions are provided in Table 11.

***DISCUSSION***

Peru is facing one of the fastest-growing equity gaps among LMICs in the distribution of the benefits of development. The isolated and mostly rural Huánuco Region on the eastern slope of the Andes mountains has one of the highest rates of child stunting (low height-for-age) in the country. Stunting reflects a child’s overall well-being. It accurately indicates social inequalities, which are the cumulative result of poor fetal growth, inadequate nutrition, and infectious disease in the first 2 years of life, associated with deficient home practices for maternal nutrition, breastfeeding, complementary feeding and micronutrient consumption, poor access to health services, and possibly also WASH practices.57

Strikingly absent in Peru and elsewhere is a system to reach vulnerable lesser-educated mothers with communication strategies that effectively change health knowledge and practices that affect maternal and child health. In the current study, we tested the effect of Sharing Histories (SH), an innovative CHW teaching methodology that is intended to improve CHWs’ cultural competencies to support improved maternal health behaviors.38

Culturally traditional women, even those with primary or secondary education, generally have special needs for effective learning. Through sharing their own experiences and hearing those of other women, CHWs-in-training become quickly attuned to a topic and become interested in hearing about practical solutions to use in specific circumstances, so they know what to suggest to prevent or solve problems. When CHWs take ownership of their past experiences and improve their abilities to express themselves, their self-esteem, self-confidence, and empowerment increase, making them more convincing and effective change agents for health behaviors and for mobilization of appropriate demand for health services.

Many countries have attempted to develop links between PHC services, community-based health care resources, and households through work with CHWs. These programs provide a channel for reaching families with information, resources, and referrals. CHWs are particularly well positioned to address health behaviors in a culturally appropriate way, but a challenge is how to help CHWs develop the cultural and personal competencies to support change in other women’s health belief systems. Government PHC staff who are CHW trainers and who use the SH
<table>
<thead>
<tr>
<th>Topic</th>
<th>Local Culturally Determined Knowledge, Beliefs, and/or Solutions Identified From CHWs Through Sharing Histories</th>
<th>Standard Messages Taught to CHWs by PHC Staff Without Sharing Histories</th>
<th>Messages Given by PHC Staff Who Are Trained to Facilitate Sharing Histories and Learning With Picture Cards</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pregnancy</td>
<td>Danger signs in pregnancy were not recognized as such. For example, female CHWs did not know that a mother could die if she is bleeding during pregnancy and may consider such bleeding as “normal.” Did not identify pain and frequency of urination as a problem. Heavy work or lifting is continued as normal. Pregnant women eat less to have smaller baby and easier birth. Many foods are specifically avoided during pregnancy, such as fish which may “impede healing.”</td>
<td>Danger signs not generally taught in a way to ensure understanding.</td>
<td>Picture cards are discussed with motivational stories of pregnant women with danger signs and how they can end in death, or how they can end well if care is sought. Mothers should spare their energy by working less eating more so the infant can have more energy. Picture cards with various images of danger signs are discussed with indications to seek care.</td>
</tr>
<tr>
<td>Birth</td>
<td>Institutional birth is not considered desirable due to fears of male health providers and horizontal birth. Women are terrorized by the idea of an episiotomy or cesarean delivery that requires transfer to a hospital distant from home and family. Care by a traditional midwife in the presence of family members is valued. Distance is a major barrier at night and holidays when no means of transport are available.</td>
<td>Home births are illegal. Institutional births are obligatory.</td>
<td>CHWs need to help mothers seek institutional birth with support from family and community members for transport.</td>
</tr>
<tr>
<td>Newborns</td>
<td>Newborn danger signs that were recognized as potentially fatal were infant not wanting to eat and infant being flaccid or agitated. Danger signs that were not recognized as such included an odor- ous umbilical stump. Newborn is placed to one side to first attend the mother immediately after a home birth, sometimes uncovered due to simple negligence.</td>
<td>Information on birth and newborns is not discussed with CHW or mothers: CHWs and mothers “don’t need to know”. Only professional birth and checkups are allowed.</td>
<td>Need for immediate drying and wrapping of newborn and placement with mother for warmth and immediate suckling at the breast. No bath the first day to stay warm. Picture card images of danger signs are provided and discussed with indications to seek care.</td>
</tr>
<tr>
<td>Breastfeeding</td>
<td>Insufficient breastmilk is a family trait, so a mother will expect it and nothing can be done if female family members had little milk. Breastmilk is withheld to avoid harming the infant if mother is angry, ill, or is pregnant again. Herbal tea is given frequently (for colic and infant thirst). Dozens of myths surrounding breastfeeding practice are expressed.</td>
<td>Generally, PHC staff are not trained in local breastfeeding beliefs and practice or in correct breastfeeding techniques. Infants are taken away for immediate newborn care and not returned quickly to the mother. Free formula samples are handed out.</td>
<td>All mothers can breastfeed if measures are taken to stimulate milk supply. Herbal tea should not be given to infants, rather the mother should drink the tea. Trainers detect local myths through CHW shared histories and use those to discuss how to avoid insufficient breastmilk and maintain exclusive breastfeeding for 6 months.</td>
</tr>
</tbody>
</table>
method can learn firsthand the cultural practices of an area to inform their training, without having to rely on costly or nonexistent in-depth research on local practices.

The CHW program described here can be implemented and fully managed at a cost of about US$1.80 per capita per year that can be divided between the health budget (US$1.20) and local government (US$0.60) (see Supplement for cost data).

We hypothesized that the experimental CHW training method could lead to maternal behavior change and subsequent improvement in child growth. This hypothesis was supported by prior research showing that recall of autobiographical memory changes future behavior.38,41,58 Prior supportive research also showed that narrative communication using firsthand and secondhand stories is an effective health communication strategy for health behavior change that subsequently improves health status.59

Qualitative findings helped to explain how change occurred. The acceptance by CHWs of the SH training method could be related to the fact that experience is the traditional basis of learning. Women’s personal experiences with childbearing and child rearing are generally not valued, but SH specifically recognizes and builds on them. Rural female CHWs are often shy and feel inhibited speaking in front of people they do not know well. The SH method provides a platform for CHWs to speak aloud in a group and practice public speaking. In standard CHW training, classes frequently begin with trainers asking questions to test CHWs’ knowledge, which can be difficult for CHWs who fear being wrong. Sharing memories, on the other hand, allows a CHW to talk about her own experiences, which are neither right nor wrong and do not require recall of something previously learned. In addition, the safe workshop environment provides each CHW with the opportunity to express herself. In this process, trust and empathic bonding develop among CHWs-in-training and trainers, especially when memories are emotionally charged. This circumstance increases the likelihood that they will collectively adjust to a new way of thinking and doing things, and it promotes mutual social support to sustain new behaviors. Furthermore, a person’s self-esteem is strengthened when they understand and take

### TABLE 11. Continued

<table>
<thead>
<tr>
<th>Topic</th>
<th>Local Culturally Determined Knowledge, Beliefs, and/or Solutions Identified From CHWs Through Sharing Histories</th>
<th>Standard Messages Taught to CHWs by PHC Staff Without Sharing Histories</th>
<th>Messages Given by PHC Staff Who Are Trained to Facilitate Sharing Histories and Learning With Picture Cards</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complementary feeding</td>
<td>When a child aged over 6 months does not want to eat, mothers give only breast milk. Mothers value giving liquid soups and semiliquid foods to infants (over semisolids). Animal-source foods are acceptable to give but are not available.</td>
<td>PHC staff recommend taking child off the breast and only give solid food. If breast milk then dries up, give milk formula or cow’s milk.</td>
<td>Continue breastfeeding and try giving small amounts of mashed food more frequently during the day. Soups are mostly water, which fills the infant’s stomach and does not allow space for the food they need to grow. Add citrus juice to legumes to make them more nutritious (increase iron bioavailability) but animal-source foods should be given as much as possible.</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>Diarrhea occurs when someone looks at the child with an “evil eye.” When diapers are damp from being hung out to dry overnight, the dampness in the diaper can “enter” the child and cause diarrhea. Traditional healers “pass a cuy (guinea pig)” or “pass an egg” over the child’s body to draw out bad energy. Dirt or lack of hygiene is not associated with diarrhea.</td>
<td>PHC staff promote use of oral rehydration fluid and care-seeking for diarrhea. Use hygiene for prevention (without discussion of local beliefs on causation).</td>
<td>Dirt on hands or on prepared food can cause diarrhea in some cases, aside from other believed causes. Thus, it is best to use hygiene practices to avoid such cases (i.e., handwashing, keep animals out of the home, keep the child off the ground or dirty floor, use correct treatment for drinking water, others).</td>
</tr>
</tbody>
</table>

Abbreviations: CHW, community health worker; PHC, primary health care.
ownership of their cultural beliefs. Research on autobiographical memory suggests that a person’s negative or “low closure” experiences have more effect on future behavior than positive or “high closure” experiences because negative memories maintain a stronger emotional charge and their details are more present in memory.39 Second-hand memories, heard from others in a group, can affect future behavior similar to one’s own memories. Through first sharing histories in a group and then learning new ways to do things from a trainer, CHWs seemed to gain enthusiasm about sharing new knowledge in a similar way to convince mothers.

A potential effect modifier of our results on stunting was the discontinuation of the National Feeding Program government food supplementation program for children ages 6 months to 3 years, that occurred half-way through the 4-year study. We postulate that it could have contributed to the plateauing of stunting prevalence from baseline to endline in the control group (Figure 3). The reduction of stunting in both study groups could arguably have been greater without this effect.

We speculate that some of the significant improvement in knowledge of danger signs and reported child feeding practices in both study groups was due to the use of the same training materials (flip charts) for all CHWs. Particularly considering the illiterate mothers in our study, their significant knowledge and reported behavior improvements did not carry through to improved growth of their children. High prevalence of stunting at 40% or above at both baseline and endline among children of illiterate mothers in both study groups may be partially explained by so-called environmental enteric dysfunction.60 We surmise that stunting in this population could not be overcome by the adoption of the specific behaviors that were promoted and measured. This explanation is supported by 2 recent major trials. A cRCT in rural Zimbabwe tested interventions to reduce stunting by addressing environmental contamination either alone or in combination with infant and young child nutritional improvements. No reduction in stunting was found except when a food supplement was provided along with nutritional counselling.61 Another cRCT in Bangladesh using direct counseling combined with a mass media campaign found improved feeding practices with no improvement in growth.62

Finally, our finding that our experimentally trained CHWs tended to visit less-educated mothers more, especially those with only primary education (Tables 4 and 7), is consistent with a recent systematic review of factors contributing to equity of CHW services that suggests CHWs tend to visit more vulnerable mothers.63

Limitations
The main limitation of our study design was the small number of study clusters, which was constrained by the number of available clusters in the study area, defined as PHC facility jurisdictions. Matching and randomization of study clusters, triple-blinding, use of DID and GEE for data analysis, and adjustment for covariates were design or analysis factors that strengthened power and reliability of the study findings.64 Our study did not measure all potential predictive factors for stunting.

CONCLUSIONS
Our study suggests that the SH method for training community health workers (CHWs) was associated with reduced child stunting when mothers were literate. Regardless of educational level, many women who live in traditional societies with culturally rooted beliefs may have special needs for effective learning for maternal and child health and nutrition. These women could be reached with behavior change strategies through CHWs trained with methods such as SH. Efforts to prevent stunting with such a behavior change strategy could be prioritized in the “low-hanging fruit” of children of women with at least some education to quickly reduce global stunting rates. The poorest children have many other determinants of stunting that are more challenging but important to address.

This study extends the research on CHW program implementation. To our best knowledge, this randomized trial is the first to test the application of autobiographical memories to help women be empowered with capabilities to serve as change agents with other women in their communities. Our study may also represent an application of neurological findings on the physical consolidation of neurons by building on memory recall to strengthen educational interventions.

More effective community health promotion is needed to attain better health outcomes in LMICs. We suggest that an integrated system should focus on strengthened government services that support and sustain careful CHW selection by women in communities, methodologically sound CHW training based on autobiographical memories used in conjunction with visual teaching material, consistent community-based supportive supervision...
Cluster-Randomized Trial on Training Method for Community Health Workers

of female CHWs, and effective monitoring and evaluation. Further implementation research on best ways to strengthen these conditions for the deployment of CHWs can contribute to meeting the basic health rights of mothers and children in LMICs and to reaching the Sustainable Development Goals.

Acknowledgments: We acknowledge the Regional Health Directorate of the Huánuco Regional Government; the Huánuco health service network management center and primary health care staff; the local government mayors and staff; and local health administration committees (CLAS) of Chinchao, Churubamba, Santa María del Valle, and Umali Districts; community health workers; and community supervisors and other community partners for their close collaboration on the project Health in the Hands of Women: A Test of Teaching Methods (En Manos de Mujeres) in Huánuco, Peru. We extend special thanks to the Huánuco Regional Health Director and staff, especially Guillermo Rengifo, María Díaz, María Moreno, Marina Santamaría, and Rosario Peña. We acknowledge the important contributions of project staff in this and previous projects to develop the community health model and CHW training program: Luis Espejo, Vilma Suárez, Alejandro Vargas, Carmen Phuyo, Dennys Talená, Lurdes Cabello, Natalie García, Jonathan Bazán, Kary Terán, Eva Hachapuma, Janeth Carrillo, Irene Quiupe, Enrique Núñez, and Marisal Borda; Yesenia Calderón, and Teleférico Félix. We also acknowledge contributions by Hilary Creed, Rocío Valverde, Miguel Campos, César Cárurmo, and others not named here. Special appreciation is due to Daniel C. Taylor for his support throughout. This work is dedicated to Carl E. Taylor, who first identified that rural women become eager to learn when asked to tell their pregnancy histories.

Funding: This field project with embedded cluster randomized controlled trial research was made possible by a competitive grant to Future Generations from the United States Agency for International Development – Office of Global Health (Child Survival and Health Grants Program, Cooperative Agreement No. AID-OAA-10-00048) for the project Health in the Hands of Women: A Test of Teaching Methods in addition to support for the Carl E. Taylor Endowment for Equity and Empowerment, Future Generations University, and local government partners in Peru. The Maternal Health Task Force of EngenderHealth through a subgrant from the Bill and Melinda Gates Foundation (GMH-107-1) provided support to an earlier project in indigenous Andean villages of Cusco to adapt the sharing histories method to Peru. The views expressed in this article are the authors’ and are not an official position of the funder.

Competing interests: None declared.

REFERENCES


49. Taylor DC, Taylor CE. Just and Lasting Change: When Communities Own Their Futures. 2nd ed. The Johns Hopkins University Press; 2016.


En español

Ensayo aleatorizado por clústeres para probar Compartiendo Historias como un método de capacitación para agentes comunitarias de salud en el Perú

RESUMEN

Antecedentes: Las agentes comunitarias de salud (ACS) se están desplegando cada vez más para apoyar la adopción por parte de las madres de prácticas saludables en el hogar. Sin embargo, se sabe poco en cuanto a la mejor manera de capacitarlas para que tengan las capacidades y competencias culturales necesarias para apoyar el cambio de comportamiento en salud de las madres. Probamos nuestro nuevo método de entrenamiento de ACS, Compartiendo Historias, en el que las ACS relatan sus experiencias de gestación, nacimiento y crianza de sus propios hijos, sobre las que se construye un aprendizaje.

Métodos: Realizamos un ensayo aleatorizado en clústeres en el Perú rural en 18 clústeres apareados. Cada clúster era el ámbito de un establecimiento de salud. Todas las ACS llevaron a cabo visitas domiciliarias con mujeres embarazadas y niños de 0 a 23 meses para enseñar, monitorear las prácticas de salud y los signos de peligro, y referirlas. El resultado principal fue baja talla-por-edad menos de -2 Z scores en niños de 0 a 23 meses. Encuestas de hogares repetidas se llevaron a cabo al inicio del estudio como línea de base (606 casos) y en la evaluación final después de 4 años (606 casos).

Resultados: Las características de las madres y sus hijos fueron similares tanto en la línea de base como en la evaluación final. El análisis de diferencias-en-diferencias (DD) mostró que los cambios promedios de talla-por-edad no fueron significativamente diferentes en los clústeres experimentales frente a los de control desde la línea de base hasta la línea final (P=0.469). Sin embargo, en el subgrupo de madres que pueden leer (alfabetas), la talla-por-edad mejoró en 1.03 en la escala de Z-score en los clústeres experimentales en comparación con los clústeres de control desde la línea de base hasta la línea final (P=0.059).

Conclusión: En comparación con los métodos de entrenamiento estándar para ACS, Compartiendo Historias puede haber mejorado la eficacia de las ACS como agentes de cambio entre las madres alfabetas para reducir la desnutrición crónica infantil. La baja talla-por-edad experimentada por los hijos de madres analfabetas puede implicar determinantes del retraso en el crecimiento que no fueron abordados en este estudio.

Peer Reviewed

Received: September 24, 2019; Accepted: October 29, 2020; First published online: December 11, 2020


© Altobelli 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-19-00332
A Rapid Cost Modeling Tool for Evaluating and Improving Public Health Supply Chain Designs

Michael Krautmann,a Mariam Zameer,b Dorothy Thomas,b Nora Phillips-White,b Ana Costache,b Pascale R. Leroueil

Key Findings
- The Rapid Supply Chain Modeling (RSCM) Tool addresses a need for more rapid and flexible ways to model the cost impact of changes to a country’s supply chain design or context.
- We compared the RSCM Tool against existing cost modeling tools and found it capable of producing similar results across a wide range of countries and supply chain designs.
- The ideal user for the RSCM Tool is a technical officer familiar with Excel and supply chain concepts; the outputs can inform both technical discussions and high-level policy decisions.

Key Implications
- Health system leaders and their technical teams should consider using the RSCM Tool to streamline the beginning stages of a supply chain design initiative, particularly in the following use scenarios:
  - Generating estimates of high-level impact to inform initial advocacy efforts
  - Sustaining momentum from initial workshops by quickly addressing supply chain questions
  - Narrowing down a wide range of initial supply chain design possibilities to help policy makers more quickly focus on the highest-impact design changes

ABSTRACT
Effective and efficient health supply chains play a vital role in achieving health outcomes by ensuring supplies are available for people to access quality health services. However, supplying health commodities to service delivery points is complex and costly in many low- and middle-income countries. Thus, governments and partner organizations are often interested in understanding how to design their health supply chains more cost efficiently. Several modeling tools exist in the public and private market that can help assess supply chain efficiency and identify supply chain design improvements. These tools are generally capable of providing users with very precise cost estimates, but they often use proprietary software and require detailed data inputs. This can result in a somewhat lengthy and expensive analysis process, which may be prohibitive for many decision makers, especially in the early stages of a supply chain design process. For many use cases, such as advocacy, informing workshop and technical meetings, and narrowing down initial design options, decision makers may often be willing to trade some detail and accuracy in exchange for quicker and lower-cost analysis results. To our knowledge, there are no publicly available tools focused on generating quick, high-level estimates of the cost and efficiency of different supply chain designs.

To address this gap, we designed and tested an Excel-based Rapid Supply Chain Modeling (RSCM) Tool. Our assessment indicated that, despite requiring significantly less data, the RSCM Tool can generate cost estimates that are similar to other common analysis and modeling methods. Furthermore, to better understand how the RSCM Tool aligns with real-world processes and decision-making timelines, we used it to inform an ongoing immunization supply chain redesign in Angola. For the use cases described above we believe that the RSCM Tool addresses an important need for quicker and less expensive ways to identify more cost-efficient supply chain designs.

BACKGROUND
Supply chains are a key component of any well-functioning health system. For vaccines, medicines, and other health products to be effective at preventing and treating disease, they must be accessible to the people who need them, when and where they are needed. Health supply chains that can reliably deliver these products to the point of care are vital to ensuring access to
To our knowledge, there are no publicly available modeling tools that are flexible enough to help decision makers evaluate the cost and efficiency of different supply chain designs while also minimizing the need for data collection and specialized software skills.

quality health care and to achieving positive health outcomes.\textsuperscript{2–4}

However, in many low- and middle-income countries (LMICs), health supply chains fail to ensure consistent availability of critical health products.\textsuperscript{5} According to the most recent World Medicines Report, a typical public clinic in sub-Saharan Africa averaged only 57% availability of its required essential medicines, and nearly 25% of all LMIC patients were regularly unable to access medicines needed for treatment.\textsuperscript{6} In addition, global vaccination coverage has plateaued at 80%–85% since 2010, and supply chain inefficiencies are considered a significant driver.\textsuperscript{7,8} These same supply chains are also expensive to operate, requiring millions of dollars of annual funding to supply thousands of public health facilities throughout a given country.\textsuperscript{9,10} For these reasons, improving health supply chain efficiency and effectiveness is a key objective for donor agencies, governments, and other health care stakeholders.\textsuperscript{11–13}

One important pathway to achieving this objective is restructuring and improving a supply chain’s design (i.e., the overarching strategy for organizing a supply chain network and its human resources, technologies, and processes). Recent studies have demonstrated that improving a supply chain’s design can lead to more cost-efficient delivery and better product availability in health facilities.\textsuperscript{14–18} Thus, donor agencies like Gavi, the Vaccine Alliance, have explicitly incorporated supply chain design into their supply chain strategies,\textsuperscript{19} and country governments are prioritizing supply chain redesign and strategy development activities in their national health plans.\textsuperscript{20,21}

The task of analyzing and identifying an improved supply chain design can often be challenging for a couple of reasons. First, detailed supply chain data (e.g., operating costs, product demand, facility locations) are often unavailable or time intensive to collect. Second, many existing tools used to collect and analyze such data are intended to provide a snapshot of the current system,\textsuperscript{22} whereas a design analysis requires flexible models that can predict the impact of large-scale changes to the supply chain. Although more sophisticated supply chain modeling and optimization tools do exist for this purpose, they typically require proprietary software and specialized modeling skills and/or consultants. In total, a supply chain review and redesign process using these current methods can require at least 3–6 months and US$250,000–US$500,000, according to recent estimates.\textsuperscript{23}

The detailed and precise outputs from such tools are necessary in some circumstances, particularly at the final stages of a supply chain design process when the focus is on fine-tuning and implementing a specific plan. However, there are a broad range of other instances where the required time and cost are prohibitive. This could include situations like conducting initial advocacy for supply chain improvements, informing workshops and meeting discussions in real time, or narrowing down a wide range of improvement options in the early stages of a supply chain design process. In such cases, leaders from ministries of health and partner organization would likely trade some level of detail and accuracy in exchange for reducing the time and cost of analysis. To our knowledge, there are no publicly available modeling tools that are flexible enough to help decision makers evaluate the cost and efficiency of different supply chain designs, while also minimizing the need for data collection and specialized software skills.

In this article, we present the design and testing of a Rapid Supply Chain Modeling (RSCM) Tool aimed at addressing this gap. We describe key attributes of the RSCM Tool, validate its results against existing supply chain analyses, and explore how it can help inform a country’s supply chain redesign process.

\section*{METHODOLOGY: DESIGNING A TOOL FOR RAPID ANALYSIS USE CASES}

The RSCM Tool is a quantitative, Excel-based model designed to quickly estimate costs and basic efficiency metrics for multiple supply chain design scenarios. It requires users to input information about the design and general characteristics of a country’s current supply chain, including: supply chain network information, such as land area and number of facilities; cost parameters, such as the cost of labor, fuel, or vehicles; and storage and distribution guidelines, such as inventory levels or frequency of delivery.

Using these inputs, the RSCM Tool models key operational supply chain activities and calculates several resulting output metrics, including: annual operating cost, disaggregated by tier and supply chain function (i.e., storage, transportation, and management); expected utilization of resources like vehicle and warehouse capacity; and operational statistics like kilometers traveled or volume delivered per facility.

Within the tool, those inputs and outputs worksheets can be replicated to create multiple supply chain scenarios, which can be compared under a main dashboard.
To address the identified need for quicker and more cost-effective decision-making tools, we made several design decisions that help the RSCM Tool maintain flexibility, quick setup time, and minimal data collection needs (Table 1).

**Simplifying Modeling Assumptions**
Like most modeling tools, the RSCM Tool uses assumptions to strike a desired balance between simplicity and accuracy. Since our goal is to provide faster results by reducing overall data requirements, we ask the user to define a typical facility at each supply chain tier, rather than requiring detailed demand, location, and cost data for every facility.

While the results do not provide detailed outputs for individual facilities, we hypothesize that for high-level, system-wide design analyses, the RSCM Tool’s outputs will be reasonably similar to those of other common supply chain modeling tools. We test this hypothesis and quantify the effect of the tool’s assumptions in the validation section below.

**Standardized “Menu” of Design Levers**
Most global health supply chains can be broken down into a relatively small set of “building block” design decisions in key functional areas like storage, transportation, and management. By incorporating these design levers along with common pre-set choices/values, a user can easily create and toggle between different distribution strategies for their supply chain network.

**Proxy Data and Worksheets to Address Data Gaps**
Since supply chain and cost data are often scarce, we incorporated several supporting worksheets and proxy datasets to help users quickly estimate

---

**TABLE 1. Key Modeling Tool Design Decisions for Facilitating Rapid Supply Chain Analyses**

<table>
<thead>
<tr>
<th>Assumptions:</th>
<th>Simplifying modeling assumptions: Reducing data requirements and enable real-time calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>• All facilities at a given tier have the same demand quantity per order period</td>
<td></td>
</tr>
<tr>
<td>• Demand is the same for every order period and does not vary over time</td>
<td></td>
</tr>
<tr>
<td>• Facilities within a tier are evenly distributed throughout a given region and, thus, are the same average distance to their nearest re-supply point</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Assumptions:</th>
<th>Standardizing design levers: Providing flexibility to model diverse global health distribution strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Storage: At which levels do you hold and manage inventory? How much safety stock does each level hold, and how frequently is it replenished?</td>
<td></td>
</tr>
<tr>
<td>• Transportation: What types of vehicles are used to transport replenishment shipments? What type of distribution model is followed at each level (e.g., hub and spoke or multi-stop distribution loops), and are there any travel constraints (e.g., administrative boundaries)?</td>
<td></td>
</tr>
<tr>
<td>• Management: Who is responsible for performing key ordering, transport, and storage functions? What types of technology supports people at each level?</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Assumptions:</th>
<th>Proxying data and worksheets to fill gaps: Enabling quick estimation of missing data points</th>
</tr>
</thead>
<tbody>
<tr>
<td>• A model for estimating immunization and/or reproductive health demand volumes and product value, by combining available demand planning methodologies with publicly available demographic and product data</td>
<td></td>
</tr>
<tr>
<td>• A general model for converting the number of units of a health product into a cubic-meter volume using historical product unit volume data</td>
<td></td>
</tr>
<tr>
<td>• Common commercial heuristics for estimating storage capacity of a warehouse based on its overall dimensions</td>
<td></td>
</tr>
<tr>
<td>• A database of typical costs for assets like vehicles, warehousing space, and cold chain equipment</td>
<td></td>
</tr>
</tbody>
</table>

Using Excel-based platform for broad accessibility
values for common data gaps, minimizing time required for data collection. Additionally, we compiled a set of complete input data templates, which are proxy data from existing cost analyses that are formatted to match the RSCM Tool’s structure. Instead of entering each input value individually, a user can “load” a preset template as a starting point for analysis, selectively overriding the proxy dataset where better data exists.

**Excel-Based Platform**

Finally, we developed the RSCM tool in Microsoft Excel because it is the most widespread software that can meet the tool’s technical requirements and run easily on most computers. Government staff and implementing partners are often familiar with and comfortable using Excel in their daily work, such as for demand planning. For users who already have Excel, there is no additional cost to accessing the RSCM Tool. Additionally, the tool is functional offline, which is essential for areas with unreliable internet connectivity. This enables the tool to be easily and widely accessible to multiple stakeholders throughout a country, which would be less likely if it required proprietary software or license fees.

### VALIDATING THE RSCM TOOL’S METHODOLOGY

To test the validity of the methodology and assumptions described, we conducted an assessment to determine whether the outputs generated by the RSCM Tool were consistent with other established methods for measuring or estimating the costs of different supply chain designs.

**Validation Approach**

Our general approach was to compile detailed datasets from recent supply chain costing and modeling analyses and replicate each analysis using the RSCM Tool. First, we built complete sets of data inputs for the RSCM Tool, compiling them from a variety of sources and vetting assumptions externally wherever possible. (The Discussion details the main challenges we faced in building these input datasets and how we addressed them.) Then, we compared the RSCM Tool’s cost estimates to the results of the original analyses.

With identical data inputs, we would expect any discrepancy in results to be driven by differences in the modeling approach and assumptions. For each comparison, we treated the existing analysis as a “reference” value and measured the RSCM Tool’s deviation from that value as an absolute percent error.

We performed this calculation at 3 levels of cost aggregation: (1) individual cost line items (e.g., fuel costs for transportation at health facilities); (2) total cost for each supply chain tier and cost category; and (3) total annual cost for the entire system. For each level of aggregation, we averaged individual error calculations together to obtain a mean absolute percent error (MAPE). A lower MAPE value implies a smaller difference between the reference and RSCM tool results.

We were able to obtain reference datasets from 6 recent cost and modeling analyses, covering 7 supply chain design scenarios (Table 2). We chose these reference analyses in part based on our ability to access underlying data since replicating the analyses as closely as possible required a more detailed breakdown of inputs and results (e.g., worker salary assumptions, specific vehicle types) than what is typically available in public reports. Additionally, we sought out analyses that were produced and vetted by country governments and partners and actually used to inform key stakeholder decisions. Even though these analyses also represent estimates of true supply chain costs, they are the best-established estimates available, and thus, serve as ideal reference values when validating the RSCM Tool.

Collectively, these analyses encompass a diverse set of current public health supply chains. They incorporate several health program areas and span a range of geographies across Africa and Latin America. They also cover several common supply chain designs, including ad hoc facility collection, “level-skipping” or “direct delivery” designs that bypass an administrative tier and a “mobile warehouse” design where facility inventory is periodically topped-up by visiting resupply vehicles.

**Adjusting Data to Ensure Equivalent Comparisons**

These 6 reference analyses were conducted by different organizations for different purposes; hence, they differ in methodological details like the scope of costs included, how costs are classified, and analysis method (e.g., simulation modeling vs. direct cost measurement at a sample of facilities). None of these methods is inherently better than another; each uses a set of data and assumptions that are tailored to its own unique context. However, due to these differences, we often needed to transform certain data inputs and outputs to
ensure an equivalent comparison with the RSCM Tool.

Many of the input parameters required by the RSCM Tool lacked a directly comparable value in the reference analyses, requiring us to make several types of estimates and adjustments, including:

- **Aligning Level of Detail:** Many RSCM inputs were available in the reference datasets but were scoped or grouped differently. For example, the RSCM Tool handles vehicle costs like fuel, maintenance, and insurance individually, but some datasets use only an aggregate “total operating cost” rate, requiring us to estimate the breakdown of that rate into its subcomponents.

- **Inferring Input Values From Results:** With some datasets (especially ones that only had results available), we lacked explicit assumptions for required inputs like vehicle maintenance cost rates. However, in many cases, we were able to infer a value from data contained in the results, such as overall maintenance costs and distances traveled. While we would ordinarily avoid using the detailed reference datasets as sources for RSCM Tool inputs, we were comfortable doing so in situations where the input parameter: (1) was an objective, numeric value, and; (2) would likely be found elsewhere in ministry or partner financial records that would be accessible in a country-level application of the tool.

- **Identifying Proxies for Missing Data Inputs:** Some RSCM inputs simply were not available in a reference dataset, often because of a difference in methodology. For example, the RSCM Tool uses a road network circuity factor to help estimate distances between facilities. We often had to use Google Maps to develop a rough proxy for this parameter because many of the reference analyses measured actual distances between sample facilities.

Similarly, when comparing final outputs, in the following examples, we often had to adjust for differences among the reference analyses in how specific cost line items were calculated. For example:

- **Costing Unutilized Assets:** Some reference analyses and the RSCM Tool track all assets that are owned by a supply chain (e.g., vehicles or storage space), while other analyses track only the fraction of those assets that are actively used. Both approaches are valid but result in different answers unless the assets are fully utilized. Thus, when comparing against this alternate approach, we scaled down the quantity of vehicles and storage in the RSCM Tool to eliminate any expected idle capacity.

- **Assigning “Ownership” of Costs:** The RSCM Tool assigns the cost of a supply chain activity to the location where that activity occurred. Certain analyses, however, assign costs to

<table>
<thead>
<tr>
<th>Study Location</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bolivia and Guatemala, 2018</td>
<td>Three supply chain costing studies led by ForoLAC (Foro Latinoamericano y del Caribe para el Aseguramiento de Insumos de Salud Reproductiva) that included all major health commodities, including vaccines: • Tarija Department, Bolivia • Quiché Department, Guatemala • Alta Verapaz Department, Guatemala</td>
</tr>
<tr>
<td>Mozambique, 2015</td>
<td>Modeling analysis conducted by VillageReach for the national and provincial ministries of health, using the HERMES software platform (Highly Extensible Resource for Modeling Event-Driven Supply Chains) to assess 2 immunization supply chain design options for Manica Province in Mozambique (the baseline 4-tier design, and a direct delivery design that skipped 1 of the tiers).</td>
</tr>
<tr>
<td>Senegal, 2017</td>
<td>Modeling analysis in Senegal estimating the nationwide costs of operating the Informed Push Model strategy for delivering family planning and maternal-child health products.</td>
</tr>
<tr>
<td>Zimbabwe, 2015</td>
<td>Evaluation of the Zimbabwe Assisted Pull System strategy in Manicaland Province in Zimbabwe, which integrated commodity distribution for most health program areas (except vaccines).</td>
</tr>
</tbody>
</table>
wherever the budget line for those costs is located. Those are not always the same locations (e.g., if health facility vehicle maintenance is funded out of a district budget).

- **Scoping Specific Cost Categories:** Reference analyses differed in the scope of costs that they were willing to consider. For example, several analyses omitted depreciation costs for health facility storage space, since the buildings are often owned by the government and require no rent or mortgage payment. Others chose to include a nominal storage space cost, since eventually that health facility building would need to be replaced.

**Validation Results**

Figure 1 shows the difference, measured in MAPE, between the RSCM Tool’s cost estimates and those of the 7 reference supply chain scenarios. MAPE values were lowest when comparing total supply chain operating costs between the RSCM estimates and the reference analyses. At this level, the only comparison to exceed a 4% MAPE value was the Mozambique baseline (6.7%). This implies that the modeling assumptions and simplifications described above generally have the smallest impact on high-level cost estimates. As the comparisons became more granular (e.g., comparing costs for an individual tier or cost line item), the differences became somewhat more pronounced.

We also saw a wide range of MAPE values across the different reference analyses. Some, such as the 2 Mozambique scenarios, saw relatively high MAPE values across all 3 levels of comparison. Others, such as Zimbabwe and the ForoLAC (Foro Latinoamericano y del Caribe para el Aseguramiento de Insumos de Salud Reproductiva) studies, were relatively low across the board. These differences appear to roughly correlate with the level of detail available in the underlying reference datasets. We discuss further implications of these results in the Discussion section.

**TESTING THE TOOL IN AN IMMUNIZATION SUPPLY CHAIN CONTEXT**

We also wanted to test the usability of the tool to understand how it aligned with real-world processes and timelines for conducting supply chain design analyses. In January 2019, United Nations Children’s Fund (UNICEF), with the support of technical partners, worked with Angola’s Expanded Program on Immunization (EPI) to explore a review and redesign of its immunization supply chain (ISC). This engagement provided us with an opportunity to evaluate whether the RSCM Tool could be used quickly and easily to produce high-level estimates during a real-time supply chain redesign.

**FIGURE 1. Comparison of Cost Estimates Between the Rapid Supply Chain Modeling Tool and Existing Reference Supply Chain Analyses**

Abbreviation: MAPE, mean absolute percent error.
Angola Immunization Supply Chain Context

Angola’s immunization supply chain consists of 4 tiers (national, province, municipality, and health facility) that align with the Ministry of Health’s administrative structure. Vaccines travel through national, provincial, and municipality stores on their way to 1,321 health facilities. By conducting a supply chain design review, UNICEF and EPI aimed to improve the efficiency of this ISC structure and understand how those improvements would impact deployment of resources like vehicles and cold chain equipment. UNICEF’s System Design Approach (Figure 2) provided an overarching framework for this design review process.

We incorporated the RSCM Tool into the data collection, validation, and modeling stage of this process after an initial stakeholder workshop in May 2019 and conducted a preliminary modeling analysis while awaiting development of a more detailed optimization model. By providing quick interim results, our goals were to sustain stakeholder interest and momentum post-workshop, get government buy-in for the scenarios defined at the workshop, and streamline subsequent modeling analyses to focus on the most promising areas of improvement.

Analysis of 4 Design Scenarios Using the RSCM Tool

We used the RSCM Tool to analyze the following 4 supply chain design scenarios that EPI representatives and partners identified at the initial stakeholder workshop.

1. A baseline scenario represented the current-state supply chain design, including the 4-tier structure, transportation strategy, and monthly resupply frequency.
2. An “ideal baseline” scenario, which maintained the current supply chain design but increased cold storage capacity to account for current shortages, represented the “true” cost to run the current design with added cold storage.
3. A reduced resupply frequency scenario lowered transport costs by switching from monthly to 2-month resupply cycles at the municipality level.
4. A level-skipping scenario bypassed the province level, with the national warehouse delivering supplies monthly to municipalities, resulting in a 3-tiered distribution structure.

The process for collecting data and analyzing these scenarios involved several key steps: (1) interviewing national-level EPI officials about key supply chain policies, (2) compiling and entering available data into an RSCM Tool data template, (3) identifying proxy data sources to fill any remaining data gaps, and (4) identifying how to model each design scenario in the RSCM Tool. In total, this process required approximately 3 weeks’ worth of personnel time, divided across 3 people. However, that time requirement could have been reduced significantly (30%–50% in our estimation) in a scenario where everyone working on the analysis was located together in Angola and fluent in Portuguese, the official language in Angola.

RSCM Modeling Results

Total cost estimates from our analysis of these scenarios are shown in Figure 3.

For the analysis, we referenced the “ideal baseline” as 100% as it represents the “true” cost of running the supply chain; this reflects the added cost Angola would need to invest in the system regardless of any design changes. Based on the results, the level-skipping approach appears to be
the most cost-efficient option, reducing annual costs by 7% over the current baseline and 11% over the ideal baseline. This difference translates to potential savings of several hundred thousand US dollars per year. Apart from the obvious reduction in storage costs, bypassing the provincial level also lowered overall transport costs by enabling more efficient transportation routes from national level to municipalities.

We presented these initial results to stakeholders from UNICEF’s Supply Division and the Angola country office in August 2019, and they will be using the results to get government buy-in for subsequent in-depth modeling. Testing the RSCM in Angola provided an opportunity to assess the RSCM Tool’s ability to quickly estimate results for stakeholders and provide guidance on which options are worth exploring in more depth.

**DISCUSSION**

We tested the RSCM Tool in 2 ways: (1) by validating it against existing supply chain costing analyses, and (2) by using it to help inform an ongoing supply chain design review in Angola. We discuss the outcomes of those testing processes and what they mean for the usability and limitations of the RSCM Tool.

**Interpretation of the Validation Results**

The validation comparisons shown in Figure 1 generally align with our expectations of what a high-level, rapid tool should be able to achieve. The RSCM Tool was very good at replicating total operating cost numbers from the reference analyses. Even at the tier- and category-specific level, estimates were often within ±5% of the reference analysis results.

It is not surprising that the RSCM Tool and the reference analyses differ somewhat in their results, given the differences in their underlying modeling assumptions. For example, we assume identical demand and travel distances for all facilities within a given tier, but there is often variation among real-world facilities (e.g., large, accessible urban health facilities vs. small, remote rural health facilities). Those differences tend to average out over a large sample of facilities, but even at a national level, this assumption likely contributes to the MAPE values in Figure 1. The data adjustments we described in the Methodology section also likely contribute to these differences.

Our key question, then, is determining what constitutes an acceptable MAPE level for the comparisons shown in Figure 1. This is challenging because the definition of “acceptable” varies with the urgency and importance of the use case. Users who need answers very quickly or cheaply are likely to accept larger discrepancies than those who have more resources or a larger decision at stake. For this question, the literature on forecasting accuracy (where MAPE is an important metric) may provide the best guidance regarding what is generally considered acceptable. Landscape reviews of published forecasts provide numerous examples of both public and private organizations willing...
to make strategic decisions with error rates of 10%–20% or more.32–34 Thus, we assume that similar rates for the validation study ( <10% MAPE at the total cost level, and <20% MAPE for individual line items), would be generally acceptable given the “rapid” use cases we are targeting. We presented this proposed threshold at several global health stakeholder consultations, workshops, and conferences, and received general approval from participants.35,36 However, as we suggest earlier, the user must ultimately define what level of accuracy is warranted.

Furthermore, like the RSCM Tool, the reference analyses from our validation are also estimates of true supply chain costs, based on other modeling tools and data analyses. Thus, they are also subject to similar accuracy challenges, including sample and observation bias, unreliable informants, and assumptions about scaling and interpreting data. We used these existing analyses as reference values because they represent the best data currently available, but they are likely not a perfect representation of true supply chain costs, and the MAPE values shown in Figure 1 should be interpreted with that in mind. Rather than being a true “accuracy” measure, the MAPE values represent evidence that, given similar inputs, the RSCM Tool can generate supply chain cost estimates that are comparable to other established methods.

**Lessons From an Immunization Supply Chain Design Context**

Although Angola’s iSC system redesign is still in process, the RSCM Tool was able to generate helpful information for stakeholders:

- Overall estimated operating cost of the current supply chain
- The rough magnitude of cost savings that could result from implementing an improved iSC design
- The estimated cost of addressing current cold chain equipment shortages
- Evidence for the types of improvements (i.e., level skipping) that will likely have the largest financial impact

Having these data points early in the design process can help the EPI better prioritize the effort it puts into ISC design improvements (and compare against other supply chain or health system improvements). It can also inform specific requests for subsequent deep-dive analyses, such as identifying optimal modes of transport for a level-skipping solution or fine-tuning the regions where this solution should apply.

The RSCM Tool generally fared well with the data and modeling challenges we encountered during this analysis in Angola. Data visibility was an issue at municipality and health facility levels, but national-level government staff were able to describe the “typical” municipality and health facility easily. We were able to use proxy data and supporting worksheets to estimate other missing values (e.g., typical demand per facility, government salary rates, and average distances between storage facilities at each tier). We also developed an estimation method to capture the effect of supplemental immunization activities, which occur in Angola every 6 months and create a temporary 30% increase in throughput during those months. The RSCM Tool does not inherently consider month-to-month differences in demand volume, but we were able to use the tool’s storage/transport utilization estimates to add buffer capacity to handle those temporary demand spikes. In these ways, we were able to address substantial data visibility and modeling challenges and produce results in a relatively quick timeframe.

Most importantly, the process of working with external stakeholders to conduct this analysis helped clarify 2 key lessons for enabling long-term external use of the RSCM Tool.

1. **Users Must Be Familiar With Data Analysis and Supply Chain Concepts**
   The tool, while designed for ease of use, is not a substitute for data analysis and supply chain knowledge. We have included various features to maximize usability of the tool (e.g., user guides, detailed interpretation notes, formatting to highlight errors and omissions), but we cannot predict all the possible analysis situations and challenges that future users might encounter. The user must be capable of making informed judgments about how to utilize imperfect data sources and align them with the RSCM tool’s structure, and how to handle novel modeling situations. Thus, a technical or logistics officer or someone familiar with assessing supply chain operating costs is likely to be the ideal long-term user of the RSCM Tool within a government or partner organization.

2. **Adopting the Tool May Require Sustained Engagement**
   Transferring long-term use of the RSCM Tool to an external organization may still require sustained engagement, at least while the tool is relatively
new and unknown. Our Angola analysis succeeded in generating interest in the tool, but it was a proof-of-concept for many stakeholders. They needed to see that initial analysis conducted before they would consider using the tool themselves the next time around. Thus, while we designed the tool so that anyone with Excel can download it and understand how to use it, we recognize that in many cases effective dissemination will involve direct engagement with country and partner stakeholders over the course of multiple use cases and analyses.

**Limitations of the RSCM Tool Methodology**

Although the RSCM Tool can cover a wide range of supply chain designs and contexts, in the following situations, the tool’s structure and assumptions are likely to create bias or error in the results.

In some supply chains a single “typical” facility may be difficult to define. This can occur for regions with unusual geographies, or with facilities highly concentrated in 1 area. It could also include supply chains with extremely variable demand over time or across facilities. These supply chains likely require more detailed data to model in the RSCM Tool, since built-in worksheets and proxy data may be less representative, and assumptions may change significantly between different scenarios.

Some supply chains utilize different strategies for different subregions or program areas. If supply chain designs are not consistent across products or facilities (e.g., a region that uses a different distribution strategy than the rest of the country), the RSCM Tool requires the user to create separate models for each design and then aggregate the results.

For small supply chain networks, a single outlier facility can greatly influence results. While the RSCM Tool can model very small regions (e.g., a province/region (generally more than 100 health facilities), there is a much greater risk that a single outlier facility could skew the results. Those outliers will tend to average out over a large enough sample, so our analyses thus far have typically modeled at least an entire province/region (generally more than 100 health facilities).

These limitations are an important consideration when deciding whether to use the RSCM Tool for a given analysis. If addressing the limitation is serious enough that it requires a large quantity of individual health facility data to resolve, then the tool may not be suitable. However, if the limitation can be addressed with a relatively small fix (e.g., manually adjusting a few parameters or creating a few scenarios), then the RSCM Tool may still be suitable.

**Broadening Access to Supply Chain Design Analysis**

Given the RSCM Tool’s characteristics and limitations, we can envision it being useful in a wide range of situations where health system leaders need quick, high-level supply chain cost and design insights but lack the resources or time for a comprehensive data collection and in-depth modeling analysis.

**Advocating for Supply Chain Initiatives**

Building initial political support for supply chain improvements often requires advocates to demonstrate the potential cost and impact of those improvements but until that support exists, resources are often unavailable to conduct in-depth and costly analysis. Being able to quickly generate high-level, country-specific data can help advocates more effectively build initial political support.

**Prioritizing Health Investment Decisions**

Health system leaders must allocate funding across numerous initiatives to maximize health impact but cannot intensively analyze all potential options. High-level supply chain cost data can help leaders better compare supply chain improvements with other health investments.

**Streamlining Traditional Supply Chain Redesign Processes**

As our engagement in Angola demonstrated, the RSCM Tool can serve as a preliminary filter for more in-depth modeling analysis, quickly narrowing down a wide range of initial design possibilities. In this way the RSCM Tool can enable a more targeted use of complex, expensive modeling software by allowing to tools to focus only on the options with the most potential. Sequencing both tools in this fashion can facilitate a quicker and lower-cost supply chain redesign process that still yields sufficient detail where needed (e.g., for final budgeting and implementation planning).

**Informing Real-Time Workshop/Meeting Discussions**

When starting from a pre-existing data template, the RSCM Tool can generate scenario analyses in a matter of minutes or hours. This opens the possibility of conducting supply chain analyses in
everyday government/donor meetings or addressing workshop questions in real-time.

**Tailoring Funding Allocations**

Many donor and government organizations allocate supply chain funding as a percentage of commodity value,22 which is a quick method but often uncorrelated with logistics costs. The RSCM Tool can provide quick estimates based on a more predictive attribute (volume) and tailored to a specific country context.

**Validating Logistics Company Bids**

Outsourcing supply chain activities is another situation where cost information is valuable but does not warrant its own extensive study. Having a readily available approximation of supply chain costs can help donors and governments more effectively negotiate with private logistics companies when they are bidding for services.

The use scenarios described in this article are not well-served by the existing landscape of supply chain tools, as they are typically diffuse, short-lived, and often too small or early stage to warrant a significant resource investment for analysis on their own. However, they are still important in building towards and sustaining supply chain outcomes. By providing stakeholders with rapid, data-driven insights in these types of diffuse situations, we can better initiate and sustain policy discussions about supply chain, more efficiently build consensus around the right types of solutions, and more effectively generate political will for larger-scale supply chain analyses and initiatives.

**CONCLUSION**

Identifying, costing, analyzing supply chain design improvements has traditionally been a highly time- and resource-intensive process. The RSCM Tool reduces these barriers, foregoing some degree of detail and accuracy to minimize data collection and enable quick turnaround of results. We tested and validated the RSCM Tool and found it capable of replicating high-level results of more traditional costing and modeling approaches. It also adapted well to existing country-level supply chain redesign processes, helping generate quick preliminary results that guide more in-depth modeling and decision making. We believe the RSCM Tool can help health system leaders make more timely and informed supply chain decisions, helping ensure efficient and reliable access to health products that are critical to improving health outcomes.

**Acknowledgments:** The authors would like to thank participants at the 2018 Global Health Supply Chain Summit and 2019 Reproductive Health Supplies Coalition, as well as colleagues from the United States Agency for International Development, John Snow, Inc., Global Health Supply Chain Program, United Nations Children’s Fund (UNICEF) Supply Division, UNICEF Angola, Pharmaceutical Systems Africa, and Merck for Mothers for their valuable contributions in shaping the scope and direction of this project. This work would not have been possible without financial support from the Bill & Melinda Gates Foundation, Gavi, the Vaccine Alliance, and the Reproductive Health Supplies Coalition.

**Competing interests:** None declared.

**REFERENCES**


Rapid Modeling Tool for Evaluating Health Supply Chain Designs


32. EOOpen. 2018 Forecasting and Inventory BenchMark Study. EOOpen; 2018.


Peer Reviewed

Received: May 10, 2020; Accepted: October 22, 2020; First published online: November 20, 2020


© Krautmann 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-20-00227
Lessons Learned From Implementing Prospective, Multicountry Mixed-Methods Evaluations for Gavi and the Global Fund

Emily Carnahan, a Nikki Gurley, a Gilbert Asiimwe, b Baltazar Chilundo, c Herbert C. Duber, d,e Adama Faye, f Carol Kamya, b Godefroid Mpanya, g Shakilah Nagasha, b David Phillips, d Nicole Salisbury, a Jessica Shearer, a Katharine Shelley, a for the Gavi Full Country Evaluations Consortium; and Global Fund Prospective Country Evaluation Consortium

Key Findings

- We present 5 key lessons distilled from 7 years of experience implementing evaluations in 7 countries, which include the importance of:
  1. Including an inception phase to engage stakeholders and inform a relevant, useful evaluation design
  2. Aligning on the degree to which the evaluation is embedded in the program implementation
  3. Monitoring programmatic, organizational, or contextual changes and adapting the evaluation accordingly
  4. Hiring evaluators with mixed-methods expertise and using tools and approaches that facilitate mixing methods
  5. Contextualizing recommendations and clearly communicating their underlying strength of evidence

Key Implications

- Global health initiatives, particularly those funding or implementing complex interventions, should consider embedding evaluations to understand how and why the programs are working to adapt as necessary and maximize impact.
- Evaluators of complex interventions should continue to share lessons learned related to balancing stakeholder priorities, aligning on “breadth” versus “depth” of the evaluation scope and ensuring use of the evaluation findings.

ABSTRACT

Introduction: As global health programs have become increasingly complex, corresponding evaluations must be designed to assess the full complexity of these programs. Gavi and the Global Fund have commissioned 2 such evaluations to assess the full spectrum of their investments using a prospective mixed-methods approach. We aim to describe lessons learned from implementing these evaluations.

Methods: This article presents a synthesis of lessons learned based on the Gavi and Global Fund prospective mixed-methods evaluations, with each evaluation considered a case study. The lessons are based on the evaluation team’s experience from over 7 years (2013–2020) implementing these evaluations. The Centers for Disease Control and Prevention Framework for Evaluation in Public Health was used to ground the identification of lessons learned.

Results: We identified 5 lessons learned that build on existing evaluation best practices and include a mix of practical and conceptual considerations. The lessons cover the importance of (1) including an inception phase to engage stakeholders and inform a relevant, useful evaluation design; (2) aligning on the degree to which the evaluation is embedded in the program implementation; (3) monitoring programmatic, organizational, or contextual changes and adapting the evaluation accordingly; (4) hiring evaluators with mixed-methods expertise and using tools and approaches that facilitate mixing methods; and (5) contextualizing recommendations and clearly communicating their underlying strength of evidence.

Conclusion: Global health initiatives, particularly those leveraging complex interventions, should consider embedding evaluations to understand how and why the programs are working. These initiatives can learn from the lessons presented here to inform the design and implementation of such evaluations.

INTRODUCTION

Complex interventions—those composed of several interacting components, sometimes with nonlinear causal pathways—are widely used to tackle complex global health challenges. As programs and interventions have become increasingly multidimensional, corresponding evaluations must be designed to assess the full complexity of these programs. Consequently, evaluations may need to consider not only programmatic outcomes,
but also other outputs and outcomes across the system to understand how to improve programs to achieve impact. The goal of these evaluations is to understand not only what happened as a result of the program, but crucially why the change occurred. This need has resulted in an increased use of mixed methods, emergence of prospective approaches, and increased emphasis on process evaluation.3–5

Gavi, the Vaccine Alliance (Gavi) and the Global Fund to Fight AIDS, Tuberculosis, and Malaria (the Global Fund) are large multilateral organizations funding country governments and partners to implement necessarily complex interventions to improve public health. In 2018, the funding disbursement of both organizations totaled nearly $USD 4.5 billion,6,7 which is being used for packages of programs including vaccine purchasing, cold chain improvements, malaria prevention programs, HIV treatment programs, tuberculosis control programs, and general health systems support. Each organization has commissioned prospective mixed-methods evaluations to examine the implementation, outcomes, and impact of these complex interventions. We define a prospective evaluation as an approach for examining implementation processes and interventions forward in time, which has several advantages over retrospective evaluation, including deeper exploration of local context and implementation barriers and facilitators, ability to monitor phases of intervention implementation, and flexibility built into the design to incorporate emerging evaluation questions. Mixed-methods approaches are increasingly recognized as critical for health systems research in low- and middle-income country contexts,8 but definitions are numerous and varied. We draw from Ozawa and Pongpirul,8 who define these approaches as evaluations that “intentionally integrate or combine quantitative and qualitative data to maximize the strengths of each, to answer questions that are inadequately answered by one approach.”

Gavi’s Evaluation Advisory Committee, a subcommittee of the Gavi Board composed of independent evaluation advisors, commissioned the Gavi Full Country Evaluations (FCE) from 2013 to 2018. The FCE was funded by Gavi and managed by the Monitoring and Evaluation (M&E) team within the Gavi Secretariat. The FCE aimed to identify drivers of vaccine coverage and equity, with an emphasis on Gavi’s support of national immunization programs.9 The Global Fund Technical Evaluation Reference Group (TERG), an independent advisory group of the Global Fund, commissioned the Global Fund Prospective Country Evaluation (PCE) from 2017 to 2021 and provides oversight through the TERG Secretariat. Like the FCE, the PCE aims to generate evidence on how Global Fund processes and policies are enacted in real time in countries to achieve Global Fund objectives.10

PATH and the Institute for Health Metrics and Evaluation (IHME) at the University of Washington have served as the global evaluation partners (GEPs) leading a consortium of country evaluation partners (CEPs) for the FCE and PCE. These evaluations cover the full spectrum of Gavi/Global Fund support, including linkages between inputs, activities, outputs, outcomes, and impact. A variety of data sources and methods are used to triangulate evidence including resource tracking, process evaluation (document review, meeting observation, and key informant interviews), root cause analysis, social network analysis, secondary data analysis, geospatial analyses, value-for-money assessments, and impact modeling (complete methods available elsewhere11–13). The evaluations aimed to understand how Gavi/Global Fund policies and processes translate into country-level implementation to provide actionable, relevant insights to improve program implementation. Both evaluations were conducted in multiple countries to produce country-specific and cross-country synthesis findings to meet the needs of country and global stakeholders. The findings have successfully influenced Gavi/Global Fund policies and processes, and it has been suggested that these types of evaluations can be used for other global financing mechanisms or initiatives.14

Our approach has shifted over time to reflect learnings gained through implementing these evaluations since 2013.11 This article adds to the existing evaluation literature, and it expands on a complementary article from the Zambia FCE team’s perspective15 by taking a broader cross-country view of lessons learned from 2 prospective mixed-methods evaluations. We present lessons learned across the evaluation life cycle to inform the implementation of ongoing or future complex evaluations.

METHODS
To generate lessons learned, we utilized our experience conducting prospective mixed-methods evaluations as part of the Gavi FCE and the Global Fund PCE, considering each evaluation as a case study. Insights came primarily from individuals who were involved in the implementation of the evaluation, both GEPs (PATH and IHME) that oversaw the evaluations and conducted
cross-country synthesis, and CEPs that were primarily responsible for data collection, analysis, and reporting in their country. The CEPs included research organizations, academic institutions, and nonprofit organizations based in the focus countries for each evaluation (FCE: Bangladesh, Mozambique, Uganda, Zambia; PCE: Democratic Republic of the Congo, Guatemala, Senegal, Uganda).

Throughout the evaluations, GEPs and CEPs generated insights through periodic internal after-action reviews and systematic reflection sessions for adaptive management. The GEPs categorized insights according to the Framework for Evaluation in Public Health (Figure) and compared our experience-based insights with existing best practices within the framework to elucidate critical differences. This framework was chosen for its straightforward, comprehensive summary of the evaluation life cycle and its widespread use. Its key steps included engaging stakeholders; describing the program; focusing evaluation design; gathering credible evidence; justifying conclusions; and ensuring use and sharing lessons.

We report on the lessons learned that add new insights to existing best practices and are likely to be the most relevant to other teams undertaking complex prospective evaluations (Box). Other lessons learned that reinforce existing practices were omitted in the interest of space, although such omission does not mean they are not important in evaluation practice.

## RESULTS

### Lesson 1: Include an Inception Phase to Engage Stakeholders

For multistakeholder evaluations of complex interventions, the evaluation team and donors should include an inception phase to focus on stakeholder engagement and evaluation design. Support by high-ranking government officials and donor organizations during the inception phase can facilitate early stakeholder engagement.

The FCE and PCE were each designed with an inception phase of 4 and 6 months, respectively. Given the complex nature of the evaluations, the inception phases were crucial to have dedicated time for a consultative and collaborative approach to engage stakeholders in developing a comprehensive understanding of the programs to be evaluated and refining the evaluation priorities and approach. Engaging stakeholders can improve evaluation design and relevance, facilitate data collection, and increase the likelihood that evaluation findings are used.

In the inception phases, we first relied on CEPs’ knowledge of the local context, reviews of
relevant technical documents, and the organizational structure of key institutions to identify relevant stakeholders. We held face-to-face meetings with individuals or small groups of stakeholders to introduce the evaluation, which was essential to get buy-in from key government officials. These meetings were followed by half- or full-day kick-off meetings in each country with a wide range of stakeholders representing ministries of health and finance, implementing partners, technical partners, civil society, and Gavi/Global Fund. In many cases, the kick-off meetings were attended or endorsed by high-ranking stakeholders such as senior government officials who encouraged support for the evaluation. For example, the Permanent Secretaries of the Ministry of Health opened the FCE inception phase stakeholder meetings in Uganda and Zambia and the Minister of Health in Senegal presided over the opening ceremony of the PCE. Support from high-ranking officials paved the way for easier access to other government officials and partners for evaluation data collection and contributed to the sense of legitimacy of the evaluation, thereby improving the likelihood that findings would be used.

Support from funders also helped facilitate support for the evaluation. At the outset, a formal letter from the Gavi CEO was shared with country governments to endorse the FCE. During the inception phase, the Gavi M&E team and the evaluation team jointly met with key Ministry of Health personnel and other stakeholders, signaling Gavi’s support for the evaluation. The PCE did not have the same level of engagement from the Global Fund Secretariat in the inception phase, in part because the PCE was commissioned independently by the TERG. The limited early engagement from the Global Fund Secretariat resulted in early challenges for stakeholder buy-in, with downstream consequences in terms of accessing information, aligning the evaluation findings with decision making timelines at the Secretariat level, and ensuring widespread dissemination and use of synthesis findings.

Lesson 2: Align on the Approach to Embedding the Evaluation in the Program Implementation

In a prospective process evaluation, the donor and evaluation team should align on the degree to which the evaluation is embedded in the program implementation; a quasi-embedded approach can balance objectivity and learning. Expectations for program stakeholders’ engagement in the evaluation should be clearly communicated by the evaluation team.

The continuum of potential evaluation designs ranges from a purely external evaluation that is entirely independent of the program implementation to a fully embedded evaluation that is internal to the implementation team.21 Purely external approaches may be more objective, but they have limited ability to understand changing program implementation, thereby potentially limiting the usefulness of the evaluation. A more embedded approach allows for more collaboration and feedback loops between the evaluation and program teams to adapt the evaluation to shifts in context,
programs, or priorities. In evaluations of complex adaptive program implementation approaches, some degree of embeddedness to understand these shifts is appropriate. Process evaluation in particular requires collaborative and trusting relationships with stakeholders involved in the program implementation to facilitate access to information. And, as noted in lesson 1, engaging stakeholders in the evaluation can encourage uptake of evaluation findings.

The FCE and PCE took a quasi-embedded approach to preserve evaluation objectivity while collaborating closely with stakeholders to support evaluation relevance, data access, and use of findings. The quasi-embedded approach can encourage timely learning through feedback loops between the evaluation teams and programs—and messaging the evaluation in this way, as a “learning platform,” helped increase stakeholder buy-in. During the initiation of the FCE and PCE, there were concerns that country stakeholders who were the most familiar with independent outcome evaluations would be resistant to the evaluation if they felt like they were being audited. Gavi anticipated this concern and emphasized to stakeholders that the FCE was not an evaluation of country programs per se, but of Gavi’s policies and processes, which helped to increase stakeholder buy-in. For the PCE, we shifted our framing to explain the evaluation as a learning platform that could provide support to stakeholders, help answer their priority evaluation questions, and provide evidence or recommendations to improve their program implementation. Many PCE stakeholders were initially unfamiliar with or had never engaged with a prospective evaluation in practice, so the concept of a learning platform was more intuitive. This framing also facilitated buy-in by providing CEPs to maintain evaluation independence (or a perception of independence). In some contexts, CEPs joining meetings solely as observers was not acceptable—they were expected to contribute if they wanted to keep being invited—thus, they shifted into participant observers. The Zambia FCE team highlighted their approach to provide meeting notes as a way of adding value, and across all CEPs, evaluator reflexivity was used to balance independence and embeddedness.

Over time, the CEPs have become increasingly embedded in country programs, for example, being added to standing program meeting invites, which also meant they were on the email distribution list to receive meeting minutes and other key documents. While this involvement has enabled CEPs to track the unfolding processes in real time, gain access to essential documents and data sources, and share back emerging findings to improve program implementation, it has also made it challenging for CEPs to maintain evaluation independence (or a perception of independence). In some contexts, CEPs joining meetings solely as observers was not acceptable—they were expected to contribute if they wanted to keep being invited—thus, they shifted into participant observers. The Zambia FCE team highlighted their approach to provide meeting notes as a way of adding value, and across all CEPs, evaluator reflexivity was used to balance independence and embeddedness.

Although the FCE and PCE have used a quasi-embedded approach to balance objectivity and learning, there has been an ongoing tension in how to strike this balance in ensuring use of findings. Evidence uptake and knowledge translation rarely occur spontaneously and must be supported through a combination of “push,” “pull,” and “exchange” activities. As a result of the messaging of the evaluation as a learning platform and the embeddedness of evaluators, the evaluation team was perceived as being well positioned to engage stakeholders in these knowledge translation activities; however, encouraging the uptake of recommendations also risked compromising evaluation independence. To create accountability with stakeholders for acting on evaluation findings, while preserving evaluator independence, the Gavi Alliance provided an annual “management response” to the FCE findings and recommendations. The management response reported how Gavi had used each finding/recommendation.

This initial messaging of the FCE and PCE as partnerships focused on learning set the stage for a more collaborative relationship between evaluators and program implementers. The CEPs established close relationships with country stakeholders who were able to share documents and data, extend meeting invitations, and serve as key informants.
it has implications for how evaluators are received and how data are collected, as well as the adoption of findings.

**Lesson 3: Continuously Monitor Changes and Adapt the Evaluation**

In evaluations of complex interventions in which the programs, organizations, or contexts are constantly evolving, the evaluation team needs to continuously monitor changes and adapt the evaluation. The evaluation plan should be designed with enough flexibility to adjust evaluation questions and approaches to respond to changes; to support this, buy-in from the donor organization is essential.

During the inception phases, the evaluation teams directly engaged the intended users of the evaluation to inform the evaluation focus, priorities, and evaluation questions (consistent with evaluation best practices). The initial terms of reference for both the FCE and PCE provided overarching strategic evaluation questions, which the evaluation teams translated into a list of country-specific and cross-country evaluation questions responsive to the organizational context during the inception phase. However, as learning institutions, Gavi and the Global Fund frequently update processes and policies, which may affect evaluation context, objectives, and priorities throughout the course of the evaluation. (For example, Global Fund’s Operational Policy Manual undergoes numerous revisions throughout the year.) Thus, understanding the program and designing a responsive evaluation was not limited to the inception phase but required an ongoing assessment as to how the program was evolving, more consistent with developmental evaluation approaches.

The quasi-embedded approach of the CEPs facilitated program monitoring at the country level, as did collaborative relationships with the Gavi M&E team and Global Fund TERG at the global level. Table 1 summarizes the approaches taken by the FCE/PCE teams to maintain direct access to stakeholders who could provide insights on the changes to Gavi/Global Fund policies and processes. Weekly calls with the Gavi M&E team and TERG Secretariat were helpful to regularly solicit updates related to policies, processes, or strategies, and buy-in from the donor organization is critical in supporting the evaluation team to fully engage with its staff (e.g., the Gavi Secretariat and Global Fund Secretariat).

As context and priorities shifted throughout the course of the evaluation, the evaluation questions had to be updated to reflect these changes, identify emerging questions the evaluation could help to address, and ensure evaluation questions are useful. Buy-in from and engagement of the donor organization and other Secretariat staff was essential to ensure relevance of the updated evaluation questions, and the approaches summarized in Table 1 served as an opportunity to validate revised evaluation questions. This ongoing monitoring of the program context and discussion of priorities resulted in the adaptation and revision of evaluation questions, and ultimately a more flexible evaluation design. Two examples of how the PCE adapted evaluation questions based on shifts at the country level and global level are included in Table 2. Although it was necessary to design the evaluation to respond to the changing program context and priorities, we experienced pros and cons associated with designing the evaluation to encourage flexibility and adaptation over time (Table 3).

As we have strengthened relationships with stakeholders, and stakeholders have a better understanding of the scope of the evaluations, this process of adaptation has become more organic,

---

**Table 1. Approaches Taken by the Evaluation Teams to Engage With the Donor Organizations to Monitor Program Developments**

<table>
<thead>
<tr>
<th>Gavi Full Country Evaluation</th>
<th>Global Fund Prospective Country Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Weekly calls with the Gavi M&amp;E team (GEP, CEP)</td>
<td>• Weekly calls with the TERG Secretariat (GEP)</td>
</tr>
<tr>
<td>• KIIs with Secretariat staff throughout the year, with</td>
<td>• Engagement with Secretariat staff at tri-annual</td>
</tr>
<tr>
<td>a concentration of KIIs during an annual in-person</td>
<td>TERG meetings (GEP, CEP)</td>
</tr>
<tr>
<td>visit to Geneva (GEP)</td>
<td>• One-off phone calls with rotating Secretariat teams</td>
</tr>
<tr>
<td>• Semi-annual touchpoints with Gavi Senior</td>
<td>scheduled by the TERG Secretariat (GEP, CEP)</td>
</tr>
<tr>
<td>Country Managers (GEP, CEP)</td>
<td>• Semi-annual touchpoints with Global Fund</td>
</tr>
<tr>
<td></td>
<td>Country Teams (GEP, CEP)</td>
</tr>
</tbody>
</table>

Abbreviations: CEP, country evaluation partner; GEP, global evaluation partner; KII, key informant interview; M&E, monitoring and evaluation; TERG, technical evaluation reference group.
with stakeholder inputs on evaluation questions shared more proactively and ad hoc. The GEPs/CEPs have also become more adept at identifying evaluation priorities through ongoing process tracking, including areas of cross-country synthesis that are most relevant in informing changes to Gavi/Global Fund policies or processes.

**Lesson 4: Include People With Mixed-Methods Expertise on the Evaluation Team**

To successfully mix methods in a complex evaluation, evaluation teams should ideally include individuals with experience across methods, or at minimum co-locate individuals with various methods backgrounds. Tools and approaches, such as collaborative data review meetings, root cause analyses, and Tableau dashboards, can help to bridge any divide between quantitative and qualitative methods expertise.

To foster mixed-methods analysis, we learned that our teams (GEP and CEP) worked best when team members encompassed various disciplinary and methods backgrounds, were co-located, and used collaborative approaches to data interpretation and synthesis. Without conscious attention to team composition and processes, we found that mixing of methods and paradigms was difficult to achieve.

Across the consortia, a range of staffing models were represented. Some CEPs had separate quantitative modeling and process evaluation teams, while others had integrated multidisciplinary teams. Aiming for a multidisciplinary team, preferably with multiple transdisciplinary staff that had “crossover” between methods expertise proved most successful. In cases in which CEP teams were divided methodologically, co-locating team members helped to ensure more regular full team meetings to review and triangulate emerging evidence, if not full mixing of methods.

To achieve true mixing of methods and paradigms, it is necessary to have both a well-integrated team with diverse expertise, as well as established procedures and processes for dialogue and analysis. While conducting a mixed-methods evaluation has been an ongoing challenge for some teams, the evaluations have adopted tools and approaches to help bridge the gap between quantitative and qualitative approaches. Using collaborative and interactive processes is valuable

---

**TABLE 2. Examples of Changing Prospective Country Evaluation Questions Due to Shifts at the Country and Global Levels**

<table>
<thead>
<tr>
<th>Responsive to Country-Level Shift</th>
<th>Responsive to Global-Level Shift</th>
</tr>
</thead>
<tbody>
<tr>
<td>In Uganda, there was an unanticipated upsurge in malaria cases in 2019, so the Prospective Country Evaluation team added an evaluation question on whether and how Global Fund policies and structures enabled the country to respond.</td>
<td>In 2020, the Grant Portfolio Solutions team at the Global Fund requested inputs about challenges related to Global Fund monitoring and reporting processes and opportunities for improvement.</td>
</tr>
<tr>
<td>Findings indicated that several flexible aspects of the Global Fund business model, including modifications to procurement and supply chain plans, facilitated the country’s response to the malaria upsurge.</td>
<td>The Prospective Country Evaluation was able to quickly incorporate new evaluation questions into the evaluation scope and shared cross-country findings to inform the Secretariat’s revised reporting guidance.</td>
</tr>
</tbody>
</table>

**TABLE 3. Pros and Cons of a Flexible Evaluation Design**

<table>
<thead>
<tr>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Is responsive to changing stakeholder needs, thereby increasing stakeholder buy-in and the like-lihood findings will be used.</td>
<td>• Can take months to get stakeholder consensus on priorities.</td>
</tr>
<tr>
<td>• Has the ability to adjust to unanticipated implementation delays to refocus on the most timely, relevant evaluation questions.</td>
<td>• Requires carefully balancing those stakeholder inputs while remaining objective.</td>
</tr>
<tr>
<td></td>
<td>• May mean that evaluation teams are developing evaluation tools in parallel to prospectively tracking a process that has already started. This may undermine the planning required for intentional mixed methods approaches.</td>
</tr>
</tbody>
</table>
Regional stakeholders in finding and refining strengths of evidence underlying each finding and should consider engaging stakeholders in the process of refining findings and recommendations.

Evaluators should clearly communicate the strength of evidence for each finding and consider engaging stakeholders in refining findings and recommendations.

In facilitating mixed-methods analysis of the data and interpretation of findings. The PCE held joint CEP-GEP data review conference calls (approximately bimonthly) to share updated quantitative analyses, discuss data quality issues and resolutions, and identify opportunities for further triangulation with process evaluation evidence or the need for additional data collection. To further facilitate collective analysis, GEP and CEP held joint in-person analysis and report writing workshops 2 or 3 times per year, in addition to cross-country synthesis workshops at least once per year. The limitation in a more collaborative analysis process is the time and cost of engaging all evaluation partners—it is a dynamic and (potentially) nonlinear process that is best served by face-to-face interaction and may take substantive time.

In terms of tools, root cause analysis was a particularly effective analytic tool as it encouraged participants to incorporate all the available data—qualitative and quantitative—and iteratively explore hypotheses collaboratively. (Example FCE root cause analyses have been shared elsewhere.) Similarly, Tableau dashboards were a useful tool to support interpretation of quantitative data among team members with a range of quantitative data skillsets; all team members had access to the dashboards and would look at the quantitative results to generate questions for qualitative follow-up. For example, the PCE visualized quantitative data from Global Fund grant revisions to understand budgetary shifts, and then generated key informant interview questions to understand why the shifts occurred and how they were affecting implementation activities.

Finally, we also learned that mixed-methods approaches can be more intentionally incorporated by starting from the evaluation question phrasing. Over time, we shifted to evaluation questions that encouraged mixed-methods data collection and analysis, such as “whether, why, and how does X outcome occur.” For example, the Uganda FCE team asked the question: “What is the effectiveness, efficiency, and country ownership of national immunization partnerships and their contribution to program performance?” This encouraged a mixed-methods approach that included social network mapping, document review, and qualitative interviews to understand the structure and added value of the partnership working on the Gavi HPV vaccine application.

Lesson 5: Contextualize Recommendations and Clearly Communicate Strength of Evidence

In evaluating complex adaptive interventions, the heightened need for attention to feasibility and context of recommendations means evaluators should clearly communicate the strength of evidence underlying each finding and should consider engaging stakeholders in the process of refining findings and recommendations.

Existing best practices focus on enhancing credibility of conclusions by ensuring data are analyzed and systematically interpreted, findings are directly linked to evidence and informed by stakeholder standards, and resulting recommendations are contextualized and actionable. While the evaluator’s role is to justify the evaluation conclusions, engaging stakeholders in the process presents a potential opportunity to further contextualize the findings and facilitate evidence use. The FCE/PCE teams shared preliminary findings with stakeholders for review to ensure we were reporting full and accurate information. Occasionally, these reviews would motivate stakeholders to share additional evidence to be incorporated. In determining when to share emerging findings, the evaluation team must balance the opportunity to gather additional insight from stakeholder reviews with the potential risk of sharing early findings with insufficient evidence that could undermine evaluators’ credibility.

Additionally, it is important to convey the strength of evidence underlying evaluation conclusions so stakeholders trust the findings and associated recommendations. This is particularly true in a mixed-methods evaluation in which each finding relies on multiple data sources with varying quality. Moreover, in some settings stakeholders perceived findings based solely on qualitative evidence to be less rigorous than quantitative evidence. To clearly signal the strength of evidence, we developed a rubric informed by GRADE and other evidence rating systems that rated the evidence along a 4-point scale. However, while the GRADE rubric considers study design and rates randomized trials highly, our scale was limited to the types of evidence used in the FCE/PCE, so randomized trials were not feasible or fit-for-purpose. Our rubric considered the extent of triangulation between data sources and the quality of the sources. Table 4 shows the strength of evidence rating used in the PCE, and Simuyemba et al. shared the rubric used in the FCE. Each finding was published with a rating to communicate our confidence in the conclusion, accounting for data quality and triangulation.

During the FCE, the evaluation team independently generated recommendations that were shared with global and country stakeholders. The PCE has taken the same approach, but in some
PCE countries we have used annual dissemination meetings as an opportunity to iteratively refine the recommendations with stakeholders. This approach has been well received and may prove a promising practice to generate buy-in for acting on the recommendations.

**DISCUSSION**

This article presents 5 lessons distilled from over 7 years of experience (2013–2020) implementing prospective mixed-methods evaluations of Gavi and the Global Fund in 7 countries. While country settings were highly variable, our experiences had some consistency, resulting in a mix of operational and practical “how to” considerations, alongside broader considerations that are sometimes more “art than science.”

The Framework for Evaluation in Public Health was a useful tool to ground the identification of lessons learned. However, while the framework suggests a distinct, linear process for evaluation, feedback loops existed between steps in practice, and some steps (e.g., stakeholder engagement) were a focus throughout the duration of the evaluations. Our lessons spanned steps in the evaluation life cycle—and are often interrelated and mutually reinforcing—and therefore we decided against presenting lessons learned aligned to specific steps in the Framework, instead emphasizing their cross-cutting nature.

Stakeholder engagement is a key theme that weaves many lessons together. In the FCE/PCE, the inception phase was the initial touchpoint to engage stakeholders (lesson 1), but strengthening relationships between evaluators and other stakeholders was an ongoing effort. The quasi-embedded approach (lesson 2) facilitated these relationships, particularly at the country level. And strong relationships—based on shared trust, collaboration, and learning—between the evaluators and stakeholders enabled program monitoring and evaluation adaptation (lesson 3), facilitated data access to inform a mixed-methods approach (lesson 4), and led to contextualized findings and recommendations (lesson 5).

A second cross-cutting theme is the balance of objectivity and learning. In recent years, the evaluation discipline has come to embrace its role in adaptation and learning, and this has extended greater latitude for how evaluator reflexivity can allow independence coupled with learning. A spectrum of evaluation models are available, depending on the nature of interactions between program implementers and evaluators and the degree of embeddedness desired. In the FCE/PCE, the quasi-embedded evaluation approach (lesson 2) allowed for timely monitoring of the program context to understand and respond to changing program needs (lesson 3). This quasi-embeddedness also allowed evaluators to communicate the strength of findings to inform stakeholders’ action (lesson 5). Ultimately, stakeholders should consider the level of objectivity and collaboration that would make an evaluation fit-for-purpose, and let that inform the appropriate degree of embeddedness in the evaluation design; there is no one-size-fits-all model for evaluation of complex interventions.

Another key theme across many of the lessons relates to the design and focus of the evaluations. Complex interventions and evaluations of those interventions often include multiple stakeholder audiences with different evaluation priorities or goals. The inception phase (lesson 1) should help define the scope of the evaluation and bring clarity to stakeholders on what the evaluation will—and, importantly, will not—address. However, we also advocate for flexibility in the evaluation design (lesson 3) to adjust evaluation questions based on shifting context, priorities, or implementation approaches. A flexible evaluation design has pros and cons (as noted in Table 3), and this is an area of continued learning for the PCE, as is discussed in the next section.

**TABLE 4.** Global Fund Prospective Country Evaluation Strength of Evidence Rating

<table>
<thead>
<tr>
<th>Rank</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The finding is supported by multiple data sources (good triangulation) that are generally of strong quality.</td>
</tr>
<tr>
<td>2</td>
<td>The finding is supported by multiple data sources (moderate triangulation) of lesser quality, or the finding is supported by fewer data sources of higher quality.</td>
</tr>
<tr>
<td>3</td>
<td>The finding is supported by few data sources (limited triangulation) of lesser quality.</td>
</tr>
<tr>
<td>4</td>
<td>The finding is supported by very limited evidence (single source) or by incomplete or unreliable evidence. In the context of this prospective evaluation, findings with this ranking may be preliminary or emerging, with active and ongoing data collection to follow up.</td>
</tr>
</tbody>
</table>

Stakeholder engagement is a key theme that weaves many lessons together.
Continuous engaging with stakeholders has been important so they know which questions have been prioritized and what types of findings to anticipate.

We recommend that future evaluations are resourced to support knowledge translation and more timely sharing of emerging findings.

Further in the Implications section. A flexible evaluation design requires an ongoing process of aligning and realigning on the evaluation questions and scope across multiple stakeholder audiences. Overall, it has been important to continuously engage with stakeholders so they know which questions have been prioritized and what types of findings to anticipate.

Implications and Future Research
Our evaluation approach has shifted over more than 7 years of implementation. As we have refined our approach, areas still remain in which we continue to learn and further refinement is required. These include balancing stakeholder priorities, aligning on “breadth” versus “depth” of the evaluation scope, and identifying approaches to ensure use of the evaluation findings.

In terms of balancing stakeholder priorities, these multilevel, multistakeholder evaluations were designed to meet the needs of a range of country and global stakeholders. It has proven challenging to design an evaluation that balances the diverse needs of distinct groups of primary users with differential interests and power. CEPs have been more likely to prioritize evaluation questions identified by country stakeholders to be responsive to country needs. Conversely, our oversight points of contact at Gavi and Global Fund have been more likely to prioritize cross-country evaluation questions that can directly inform policies or strategies or are responsive to their funders and board members. With limited resources, if tradeoffs needed to be made between being responsive to global versus country priorities, it was not clear which to prioritize. Striking a balance between stakeholder priorities has been an ongoing challenge.

A second area of continued learning is how to align stakeholders on the tradeoffs between covering a wide breadth of topics versus going in depth on fewer topics. In setting the evaluation questions, the FCE and PCE teams have continuously navigated the tradeoffs between depth versus breadth of the evaluation scope. Process tracking (through document review, meeting observation, and key informant interviews) was intended to understand the breadth of activities, and based on stakeholder priorities and emerging findings, evaluation questions could be identified for further in-depth analysis. However, in practice, it has been challenging for CEPs to track all the processes unfolding—particularly for the PCE since it covers 3 large disease programs, with many stakeholders and grant activities. Over time, both the FCE and PCE shifted toward less breadth and more depth, with more focused evaluation questions and analytical approaches. On reflection, it was important for the evaluation teams to start with a broad scope to understand all the interrelated components of the complex interventions; with this understanding in place, it was possible to narrow the evaluation focus to go further in depth without losing the wider context.

Finally, we continue to test and refine our approach to ensuring use of the evaluation findings among target audiences. Lessons 1 and 2 highlight our approach to engaging with stakeholders, which engenders buy-in to the evaluation and uptake of findings. Best practices emphasize tailoring dissemination strategies to stakeholders and providing knowledge translation support; however, the FCE and PCE teams have had limited resources and capacity to support this effort. Our more formalized dissemination approaches have focused primarily on annual written reports and annual country-based dissemination meetings. Annual dissemination meetings have worked well to bring together a diverse set of stakeholders to discuss evaluation findings and recommendations and provide input on future evaluation priorities. However, the timing of annual meetings and reports may not align with program implementation timelines or decision-making windows. Thus, it is important to have multiple modes of disseminating findings. We recommend that future evaluations are resourced to support knowledge translation and more timely sharing of emerging findings (e.g., through shorter policy briefs, evaluation team engagement in program meetings) to fully take advantage of the learning platform.

Limitations
The content for this article draws solely from the experiences of the FCE and PCE evaluation teams, meaning the lessons do not directly incorporate the perspectives of other key stakeholders (e.g., Gavi, Global Fund, Ministries of Health) on what aspects of the evaluation worked well and added value versus those needing further refinement. Furthermore, the lessons presented are not exhaustive; the authors’ judgment was used to determine which lessons were most novel and important to highlight. Another potential limitation is that lessons are drawn only from the FCE and PCE cases, which are unique evaluations in scale and scope and not necessarily generalizable. However, the case uniqueness also suggests
lessons may be particularly relevant to other large global health initiatives with interest in establishing similar multiyear, independent prospective evaluations of their investments, policies, and processes.

**CONCLUSION**

A key benefit of prospective mixed-methods evaluations is the opportunity for dynamic and continuous learning because data are collected while implementation unfolds. This means that evaluators can identify what is working or not working and explore why. Although this type of evaluation has added value to Gavi’s and Global Fund’s understanding of their programs, this approach is a new way of working for many evaluators, donors, and other stakeholders, meaning it can take time to understand and engage with. Therefore, this article presents 5 lessons distilled from over 7 years of experience (2013–2020) implementing prospective, mixed-methods evaluations of Gavi and the Global Fund in 7 countries. Our aim in writing this article was to reflect on and share key lessons that we hope can inform the design and implementation of future prospective evaluations of large-scale, complex global health initiatives. Such global health initiatives, particularly those leveraging complex interventions, should consider embedding evaluations to understand how and why the programs are working to adapt as necessary and maximize impact.

**Acknowledgments:** The authors acknowledge the important contributions of all evaluation team members who designed and implemented the Gavi Full Country Evaluations and the Global Fund Prospective Country Evaluations from a coalition of organizations: the Institute for Health Metrics and Evaluation (IHME), University of Washington (USA); PATH (USA, DRC, and Senegal); the International Centre for Diarrhoeal Disease Research (icddr, b) (Bangladesh); Universidade Eduardo Mondlane, Health Alliance International, Manhiça Health Research Centre, and National Institute of Health (Mozambique); Infectious Diseases Research Collaboration (IDRC) (Uganda); University of Zambia (Zambia); Centre de Investigación Epidemiológica en Salud Sexual y Reproductiva (CIESAR) (Guatemala); and Institut de Santé et Développement/Université Cheikh-Anta-Diop (ISED/UCAD) (Senegal). The authors also acknowledge Gavi, the Vaccine Alliance and the Global Fund for AIDS, Tuberculosis, and Malaria for supporting the design and implementation of these evaluations. (In doing so, we do not imply they endorse this document.) We particularly thank the Gavi Monitoring and Evaluation team and the Global Fund Technical Evaluation Reference Group (TERG) and TERG Secretariat for their partnership. Finally, we extend our thanks to all stakeholders in the evaluation countries and Geneva who participated in the evaluations.

**Competing interests:** The authors received financial support from Gavi, the Vaccine Alliance and the Global Fund for AIDS, Tuberculosis, and Malaria to conduct the evaluations.

**REFERENCES**


Peer Reviewed
Received: March 30, 2020; Accepted: September 30, 2020; First published online: November 11, 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-20-00126
Effects of a Community-Based Program on Voluntary Modern Contraceptive Uptake Among Young First-Time Parents in Cross River State, Nigeria

Gwendolyn Morgan,a Anjala Kanesathasan,b Akinsewa Akiodec

Key Findings

We implemented and evaluated a program to improve child spacing, modern contraceptive use, and related gender outcomes among first-time parents in Cross River State, Nigeria.

- Contraceptive awareness, attitudes, and couples’ communication increased significantly from baseline to endline.
- After controlling for significant factors related to family planning use, first-time mothers were 3.3 times more likely and male partners 3.7 times more likely to be using a modern contraceptive method at endline. Most first-time mothers and their partners preferred the contraceptive implant, and a smaller percentage chose the injectable.

Key Implications

This experience suggests that local and state governments can adapt and scale up 3 essential program elements:

- Ensure the availability of modern contraceptive methods through local health facilities
- Use community-based health workers to conduct home visits with first-time parents to provide tailored health information and referrals, as well as build linkages with the formal health sector
- Include activities that address gender norms and couple dynamics to foster better alignment, communication, and joint action on reproductive issues

ABSTRACT

**Background:** Reproductive health programs for youth have largely overlooked first-time parents (FTP)—defined as young women younger than 25 years old who are pregnant or already have 1 child, and their partners. To address this gap, we implemented and evaluated a program to improve child spacing, modern contraceptive use, and related gender outcomes among FTPs in Cross River State (CRS), Nigeria. This paper examines the effectiveness of FTP interventions in improving voluntary uptake of contraception.

**Methods:** We conducted small group sessions and home visits with FTPs from May to August 2018 in 2 local government areas of CRS. A pretest–posttest study examined the effectiveness of these interventions regarding healthy timing and spacing of pregnancy/family planning knowledge, attitudes, intentions, communication, decision making, and contraceptive use. We performed a bivariate analysis and logistic binomial regression to confirm change over time in the primary study outcome, current use of a modern method of contraception. We also performed analysis of demographic characteristics and secondary outcomes (e.g., birth spacing intentions and couple communication).

**Results:** We interviewed 338 participating first-time mothers (FTMs) and 224 participating partners at baseline and endline. Important indicators of contraceptive awareness, attitudes, and couples’ communication increased significantly from baseline to endline. Voluntary current modern contraceptive use increased from 26% to 79% among nonpregnant FTMs ($P<.000$), and from 44% to 81% among partners ($P<.000$). After controlling for significant factors related to family planning use, FTMs were 3.3 times more likely ($P<.001$) and partners 3.7 times more likely ($P<.000$) to be using a modern contraceptive method at endline.

**Conclusion:** Program participation was associated with significant improvements in voluntary uptake of modern contraceptive methods and multiple secondary outcomes. Even within a short timeframe, this intensive, multi-intervention effort achieved significant advancements across healthy timing and spacing of pregnancy and family planning outcomes for this vulnerable youth population.

INTRODUCTION

Longer birth intervals, facilitated by modern contraceptive use, are associated with reductions in maternal and neonatal mortality and morbidity in low- and middle-income countries. Adolescent mothers around...
the world disproportionately experience pregnancy-related death and disease when they start childbearing early and have rapid repeat pregnancies. Global studies show that adolescents aged 15–19 years have less access to voluntary modern contraception, use modern contraceptives less frequently, and have a higher unmet need for modern contraception than older women.3–5 These factors place adolescent and young mothers at risk of negative health outcomes and highlight the particular vulnerabilities that young women and mothers face, including those going through pregnancy, childbirth, and childrearing for the first time.

This global pattern is reflected in Cross River State (CRS), Nigeria. As of May 2017, available national- and state-level data showed that sexual activity and motherhood began early. In CRS, 18% of adolescent girls aged 15–19 years have started childbearing.6 In addition, adolescent and young mothers often do not use modern contraception to space their second child or subsequent children. As a result, rapid repeat pregnancies are common, with nearly one-quarter of all children born less than 2 years after a sibling.7 Only 27% of sexually active adolescent girls (15–19 years) in CRS reported using a modern contraceptive method (both married and unmarried).7

As noted in a 2007 literature review by the World Health Organization8, the social and economic consequences of adolescent sexuality and pregnancy greatly depend on an adolescent’s particular cultural, family, and community setting. In CRS (as in other parts of Nigeria and in the global context), where fertility is highly valued within the institution of marriage, unmarried young mothers often endure additional stigma and discrimination within their communities and families due to an early, unplanned, and unwanted pregnancy.9–11 As a result, they may experience adverse social consequences such as curtailment of their education, decreased mobility, financial deprivation, and increased social isolation.9,12 Young men who unexpectantly become fathers for the first time may also face numerous challenges in continuing their education and providing adequate financial support for themselves and their new family.9,13

Adolescent and young mothers also face critical barriers (financial, physical access, family permission, etc.) in seeking health care and services.9 Key influencers, such as parents, in-laws, husbands or male partners, and perhaps older co-wives, typically drive household decision making as well as health care spending. Evidence shows that couples who discuss and jointly make decisions about family planning and use of contraception, as well as receive social support for using contraceptive methods and services, are more likely to use contraception.14–16 Men who approve of their female partners’ contraceptive use, provide support to obtain transport to reach a facility, and provide funds and permission to access services are critical to facilitating women’s contraceptive use in many country contexts, including Nigeria.17,18 Yet male partners, parents, and even in-laws may implicitly or explicitly discourage use of contraception due to concerns about perceived and actual side effects. They may also simply fail to give permission to visit a health facility or give financial support for these services to new young mothers, especially those who are not yet empowered to initiate conversations about family planning.

In addition, adolescents and young women themselves have their own biases and misinformation about the risks and potential side effects of contraception, such as infertility, permanent damage to reproductive organs, infections, or cancer.19,20 Methods that interrupt the perceived natural pattern of menstruation are largely deemed unacceptable. Adolescents and young women often perceive that they will be more likely to experience these side effects if they use long-acting contraception, such as an intrauterine device or implant. Improving the attitudes of adolescent and young mothers and their partners about the healthy timing and spacing of pregnancies (HTSP) and supporting their informed knowledge and voluntary use of modern contraception are particular priorities for health programs, given the higher risks of morbidity and mortality for both the mother and the child.21

First-time parents (FTPs)—defined as young women under age 25 who are pregnant or already have 1 child, and their partners—have largely been overlooked in reproductive health programs for youth. A 2014 review of global data showed that many first-time mothers (FTMs) are at increased risk of poor pregnancy, delivery, and child health outcomes, a situation compounded by multiple factors that limit their access to timely health information and services.22 The needs of FTPs extend beyond the scope of many adolescent and youth programs, which often cater to unmarried clients and focus on the prevention of pregnancy. Issues faced by young parents, such as infant care and feeding or couple communication and decision making are also not typically included in family planning and pregnancy prevention programs aimed at women of reproductive age or even married youth.
To address this gap in CRS, the Evidence to Action (E2A) Project, a global family planning project funded by the United States Agency for International Development, and Pathfinder International/Nigeria launched a program to improve child spacing, voluntary contraceptive use, and related gender outcomes among FTMs and their male partners. Implemented through the Saving Mothers, Giving Life (SMGL) Initiative, and in partnership with the CRS Ministry of Health, the program focused on increasing access to HTSP and family planning information and services, as well as addressing the underlying social and gender factors that influence family planning communication, decision making, and action for FTMs and their male partners. The program applied both a life course and a sociocultural lens to determine the appropriate content and structure of interventions with young FTMs, their husbands/partners, other key influencers, and the broader community. We also built on existing facility- and community-based family planning services, strengthened under the ongoing SMGL Initiative, to provide targeted family planning counseling and referral linkages for FTMs and their male partners. The package of FTP interventions was implemented in 2 local government areas (LGAs) of CRS, Ikom and Obubra, from May through August 2018.

This article examines the effectiveness of community-based FTP interventions in improving FTPs’ demand for HTSP and their voluntary uptake of contraception through analysis of key indicators obtained from the baseline and endline survey results among program participants. These indicators, which were part of the initial conceptual model and reflect program content on HTSP and family planning, include intentions to space the next birth by at least 3 years; awareness of 3 or more modern contraceptive methods; belief that contraception will “spoil” or harm one’s reproductive organs, perceived approval from a male partner for a female partner to use contraception, recent partner communication about contraceptive use, perceived joint decision making about using a method of contraception, and finally, current voluntary use of a modern method of contraception.

**PROGRAM DESCRIPTION**

From 2014 to 2019, E2A and Pathfinder International worked closely with the CRS Ministry of Health and other partners to decrease the maternal mortality ratio and neonatal mortality rate and increase contraceptive use across the state through the SMGL Initiative. Using a systems approach, SMGL strengthened state, facility, and community networks to address the 3 delays that contribute to maternal mortality (delays in deciding to seek appropriate services; reaching those services; and receiving timely, quality care once the service site is accessed) and increase access to comprehensive family planning services, including long-acting reversible contraceptives, in 108 facilities in all 18 LGAs of CRS. Although the SMGL Initiative achieved reductions in facility neonatal mortality rate and facility maternal mortality ratio, the project team noted a persistent gap in reaching young women and mothers with family planning services at the community level, including those at risk of early childbearing and rapid repeat pregnancies.

Informed by evidence from formative research conducted with FTMs in 2017 (Box 1), E2A designed a program to improve voluntary modern contraceptive use and related gender outcomes among FTMs and their male partners. The FTP interventions built on existing SMGL service delivery and community platforms in 2 LGAs (Ikom and Obubra), selected on the basis of local capacity to implement community-based activities and to engage FTMs, their husbands/partners, other key influencers, and the broader community to improve contraceptive access and use. These LGAs also had sufficient numbers of adolescent and young women who were potentially FTMs. While specific data on exact numbers of FTMs in these 2 LGAs are not available, the 2015 census projections estimated that the total population of these 2 LGAs was 433,363, and approximately 9% (or 39,002) were females aged 15–24 years. While the provision of modern contraceptive methods was largely done at health facilities, the FTP interventions expanded community-based activities to deliver HTSP and family planning information, as well as counseling and referral services, and to address underlying social and gender factors that influence family planning–related communication, decision making, and action. Staff from SMGL and a local community-based organization (CBO) partner, the Greater Hands Foundation, implemented activities in 16 health facilities (a subset of public sector and faith-based facilities working with SMGL in Ikom and Obubra LGAs) and 37 communities served by these facilities. Preparations for FTP activities began in early 2018, with the main period of implementation occurring from May through August 2018.

FTP interventions included peer group sessions with FTMs; small group sessions with the
modern contraceptive uptake among first-time parents in cross river state, nigeria

Evidence to action conducted formative research with first-time mothers (FTMs), male partners, mothers of FTM, and other respondents in Cross River State, Nigeria, in May 2017. The following key findings informed the design of the new first-time parent (FTP) component:

- Nearly all FTM and male partners agreed that birth spacing is beneficial for the mother, infant, and family, and could name at least one benefit of child spacing.
- Most FTM and male partners could name or describe at least one modern family planning method, but did not know how to use any of the methods.
- Some FTPs were not sure whether family planning was safe and were concerned that its use might “spoil the womb,” thereby negatively affecting a woman’s future fertility. Married FTM generally thought family planning was safe and beneficial, while unmarried FTPs (especially those that had never used contraception) were less likely to believe that family planning is safe for young mothers to space their children.
- Several men mentioned that they prefer “the local method” of spacing (extended postpartum abstinence), and a few men mentioned that family planning is only appropriate for women who have finished childbearing or for women in school so that they can “concentrate on their studies.” Despite this apprehension about the safety of family planning, most reported that they would approve of their wives/partners using family planning if they wanted to do so to space their births.

The formative research findings pointed to limited awareness and use of family planning services and a need to increase awareness across study sites. Recommendations also included provision of accurate and comprehensive information on family planning methods, providing effective counseling on family planning methods and services, and encouraging spousal communication to improve family planning decision making and uptake. The findings also noted young women/mothers’ limited use of health facilities, highlighting the need for community-based approaches that reach young people and link them to the larger health system.

Among FTP interventions, FTM peer groups, small groups with husbands/partners, and home visits by CVs were particularly important for family planning.

FTM Peer Groups

The core FTP intervention was a small group activity with FTP, grounded in the concept of creating safe spaces, peer networks, and role models for young women going through similar life experiences. Fifty groups were established in May 2018, each led by a young FTP peer leader and composed of 12–15 members. Groups met weekly in their communities for 14 sessions over the 4-month intervention period. At each 1-hour session, the peer leader used an activity card to guide discussions on a specific health or gender topic, such as HTSP, a modern contraceptive method, or problem solving within relationships. CVs generally attended all sessions to support peer linkages with facilities, monitoring reports). CVs participated in implementing all elements of the FTP component and were the linchpin between different activities, especially in connecting FTPs and communities with health facilities. Field activities were closely monitored by project and CBO staff attending project activities to observe progress, provide supportive supervision, and assist with any troubleshooting. Three FTP interventions—FTM peer groups, small groups with husbands/partners, and home visits by CVs—were particularly important for improving family planning-related knowledge, attitudes, gender dynamics, and actions and for increasing access to family planning services.

Among FTP interventions, FTM peer groups, small groups with husbands/partners, and home visits by CVs were particularly important for family planning.

hospitals/partners of peer group members; small group sessions with older women, typically the mothers or mothers-in-law of peer group members; home visits by Greater Hands Foundation community volunteers (CVs); community sensitization; and ongoing family planning service delivery at facilities and through mobile outreach. While the Greater Hands Foundation had already been active in these communities through SMGL, new FTP activities required increased CV capacity and engagement. E2A worked with the foundation to recruit 25 certified, but not yet employed, community health extension workers (CHEWs), a health worker cadre with 2–3 years of training and typically based in peripheral health facilities. Recruited CHEWs lived in FTP intervention communities and were awaiting their official Ministry of Health posting. These CHEWs (21 women and 4 men) served as CVs dedicated to reaching FTPs and received a monthly stipend (20,000 naira, or approximately US$55 at the time of the intervention) and transport allowance (20,000 naira, or approximately US$55) to conduct activities. CVs were trained by project and CBO staff on priority health issues, including danger signs and 3 delays during pregnancy, HTSP and family planning, exclusive breastfeeding, positive parenting, and gender norms and barriers to accessing services. Trainings also stressed communication and facilitation skills, as well as project roles and responsibilities (e.g.,
leaders, answer questions, and schedule home visits. In total, 599 out of 607 peer group members attended at least 12 of the 14 sessions.

**Small Group Sessions With Male Partners**

The FTP program prioritized a structured intervention with the male partners of FTM peer group members, given their influence over health decisions, including family planning use, and their own needs. By design, the male partner intervention began after the FTM peer groups, giving FTMs time to determine if they wanted to include their husband/partners. Once identified by the FTMs, CVs and “male motivators” (the partners of FTM peer group leaders) invited husbands/male partners to the small groups. This peer-to-peer approach worked well, as men were comfortable discussing the proposed activity with other men and also appreciated knowing someone who would be in the group. In total, 20 male partner groups formed in July 2018, engaging 241 men, against a target of 200, in 6 weekly sessions. These sessions explored similar health- and gender-related topics as the FTM peer groups, including HTSP/family planning information, counseling, and referral services, but also addressed other pre- or postnatal issues as relevant. CVs conducted 4–6 home visits with each peer group member from May to August 2018, often at the request of the FTM or male partner, or in follow-up to an earlier conversation or referral. As much as possible, CVs made an effort to engage male partners, older women, and other household members, and often helped to address different or conflicting perspectives on possible health actions. Home visits accounted for the majority of family planning referrals given and completed. The multiple points of contact over the 4-month intervention period proved instrumental in building FTPs’ trust and confidence in CVs and, importantly, creating linkages with the broader health system.

**METHODOLOGY**

This evaluation employed a quantitative pretest-posttest design with program participants to evaluate outcomes related to knowledge, attitudes, and behaviors on family planning and HTSP, exclusive breastfeeding, child development and parenting, and gender-equitable relationships. Both male and female CVs led these sessions, as their status as community resource persons helped overcome any inhibitions felt by male participants. Almost all men (231 of 241) attended the full set of sessions.

**Household Visits by CVs**

Under SMGL, CVs conducted home visits during pregnancy, immediately after delivery, and 6 weeks postpartum to support maternal and infant health outcomes. The FTP component supported home visits further into the extended postpartum period for FTM peer group members. CV visits focused primarily on HTSP/family planning information, counseling, and referral services, but also addressed other pre- or postnatal issues as relevant. CVs conducted 4–6 home visits with each peer group member from May to August 2018, often at the request of the FTM or male partner, or in follow-up to an earlier conversation or referral. As much as possible, CVs made an effort to engage male partners, older women, and other household members, and often helped to address different or conflicting perspectives on possible health actions. Home visits accounted for the majority of family planning referrals given and completed. The multiple points of contact over the 4-month intervention period proved instrumental in building FTPs’ trust and confidence in CVs and, importantly, creating linkages with the broader health system.

**BOX 2. Topics Addressed in First-Time Mother (FTM) Peer Groups, Cross River State Nigeria**

Evidence to Action (E2A) adopted 12 activity cards from a toolkit developed by the Gender Roles, Equality, and Transformations (GREAT) project, led by the Institute for Reproductive Health of Georgetown University and implemented by Pathfinder International and Save the Children in Northern Uganda.26 E2A developed 2 additional cards, one on exclusive breastfeeding and the other on positive parenting. Topics included the following:

- Healthy timing and spacing of pregnancy
- Problem solving in intimate relationships
- Life aspirations
- Contraceptive methods: implants, injectables, oral contraceptive pills, condoms, emergency contraception
- Gender norms
- Communication and decision making among couples
- Desired family size
- Gender-based and intimate partner violence
- Exclusive breastfeeding
- Positive parenting

The evaluation focused on knowledge, attitudes, and behaviors on family planning and HTSP, exclusive breastfeeding, child development and parenting, and gender-equitable relationships.
members/participants of intervention groups (small group sessions), before and after participation in these groups.

Sample Size
Using a sample size calculation, the program team determined that a sample of 300 FTM peer group members and 200 male partner group members would be sufficient to detect a 10-percentage-point increase in current use of family planning (a key program outcome indicator) from an assumed baseline value of or near zero. This would yield a sample detecting a significant difference from baseline to endline at the \( P < 0.05 \) level of significance with a design effect of 2.0. A 2-stage (peer groups and individual members) cluster sampling scheme was used to sample FTM respondents. The research team proportionately allocated the FTM sample (\( N = 300 \)) among each of the 2 LGAs based on the total number of participants and peer groups in each LGA. The study team randomly sampled respondents at both baseline and again at endline from the same 32 FTM peer groups. Due to the smaller size of the male partner program, a research team interviewed all male partners participating in the program from each of the 20 male partner groups at baseline and endline. The final achieved sample size was 338 FTMs at baseline, 339 FTMs at endline, 245 male partners at baseline, and 225 male partners at endline (Table 1).

Ethical Review
The study protocol and other required documents were submitted to the Government of CRS of Nigeria Health Research Ethics Committee (CRS-HREC) in Calabar, Nigeria, and to PATH’s research determination committee (RDC) in the United States in late 2017. E2A and Pathfinder International received approval to proceed with the research from the CRS-HREC on March 2, 2018. On February 26, 2018, PATH’s RDC approved the application and determined it to be “not research,” therefore obviating the need for any additional U.S.-based institutional review board review, including PATH/US institutional review board.

Data Collection
Baseline data collection took place May 9–18, 2018, for FTMs and July 9–15, 2018, for male partners, and endline data collection for both FTMs and male partners took place from August 20 to September 2, 2018. At baseline, interviews took place during the initial group activities; a trained research team conducted private, one-on-one interviews with recruited FTMs and male partners/fathers who agreed to enroll in their respective group-based activities and consented to participate in the study. At endline, participants were recruited for private one-on-one interviews at the conclusion of the final group session. The research team of field-based staff conducted face-to-face structured interviews using standardized, precoded questionnaires at both baseline and endline.

<table>
<thead>
<tr>
<th>Selected Participants</th>
<th>Baseline</th>
<th>Endline</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ikom</td>
<td>Obubra</td>
</tr>
<tr>
<td></td>
<td>Ikom</td>
<td>Obubra</td>
</tr>
<tr>
<td>First-time mothers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At least 10 FTM members were randomly sampled at both baseline and endline from each of the selected peer groups in both LGAs</td>
<td>15 of 24 peer groups randomly selected at baseline; 150 FTMs randomly selected from each of the 15 groups</td>
<td>17 of 26 peer groups randomly selected at baseline; 188 FTMs randomly selected from each of the 17 groups</td>
</tr>
<tr>
<td>Male partners of FTMs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All male partners participating in peer groups in both LGAs were selected for the study and interviewed</td>
<td>All (census) 10 peer groups selected at baseline; all members of each group interviewed at baseline; 123 male partners interviewed</td>
<td>All (census) 10 peer groups selected at baseline; all members of each group interviewed at baseline; 114 male partners interviewed</td>
</tr>
</tbody>
</table>

Abbreviations: FTM, first-time mother; LGA, local government area.

Table 1. Criteria for Selection of Respondents and Achieved Sample Size Among Young First-Time Parents, Cross River State, Nigeria
endline. For all interviews, participants received a summary of the study and were requested to sign a consent form (with provisions for thumbprint signatures). Signed consent was obtained and a copy given to participants. Interviews were conducted in either English or Pidgin language.

**Data Management and Analysis**

The research team collected data using Android-based mobile phones with the Open Data Kit application. The mobile phone data entry application included built-in consistency checks and skips. The research team uploaded the dataset to a platform storage server, where it was monitored centrally during the period of field data collection. The team then downloaded the dataset to Excel and cleaned, labeled, and checked it for inconsistencies. SPSS Version 22 was used to perform a descriptive data analysis, using simple frequencies and bivariate analyses. Based on the sampling scheme, the baseline and endline FTM samples (which were randomly generated at both times) were treated as independent samples, and the male partner samples were treated as nonindependent repeated measures. Although FTMs were randomly recruited at baseline and endline, the peer groups to which they belonged were the same at baseline and endline. Therefore, the authors of this study conducted a post hoc analysis of independence between the 2 FTM samples and determined that about 75% of the sample at baseline was included again at endline. A sensitivity test was therefore conducted with only the repeated measures subset to confirm the robustness of the analysis and the statistical significance of the FTM findings. All statistical comparisons of the FTM data presented in this paper were reanalyzed as repeated measures using McNemar’s test and paired t-test were used to present statistical differences at baseline and endline using the matched sample. In addition, all statistical comparisons of the male partner data presented in this paper were analyzed using these tests with and without complex sampling (based on 2-stage cluster sampling). Significance levels of these findings did not change based on adjusting for 2-stage cluster sampling, remaining highly significant.

### RESULTS

**Sociodemographic Characteristics**

The baseline and endline survey data provided useful information about the characteristics of FTMs and a subset of their male partners who joined and stayed engaged in interventions. While some recruitment inclusion criteria were set for FTM peer group members (under 25 years, pregnant or with first child) and their male partners (identified and nominated by interested/willing FTM participants), activities were otherwise open to FTMs and male partners who wanted to participate. Table 2 presents select background characteristics of FTMs and a nominated subset of male partners engaged in the FTP interventions. Almost all FTMs were within the required age limit at baseline, with roughly 63% aged 20–24 years and 29% aged 15–19 years. Participating male partners were most likely to be older; 30% of male partners were aged 30 years or older at baseline and endline.

At baseline, most FTMs reported that they were not married/living with their partner (63%, N=338), and 68% of a subset of nominated male partners (N=245) reported that they were either married or living with their partner. The majority of FTM participants (86%) had 1 child with a mean age of 6.9 months at baseline, with another 14% pregnant with their first child. The data also show that most (90%) male partners enrolled in the program were also first-time fathers. A majority of both FTMs and male partners who participated in the FTP program reported completing a secondary or higher level of education. Although most (85%) male partners reported being currently employed at baseline, only about one-third of
FTMs (36%) reported working at baseline, likely due to their recent pregnancy and delivery. Between baseline and endline, FTMs had a few significant differences in some of these variables. FTMs at endline were slightly older (21.1 years of age) than at baseline (20.6 years of age), as were their babies, largely due to the 4-month interval between data collection efforts for FTMs. (Only

| TABLE 2. Percentage Distribution of Age, Marital Status, Local Government Area, and Education Level by Participant Group and Baseline/Endline Among Young First-Time Parents, Cross River State, Nigeria |
|---------------------------------|-----------------|-----------------|
| Age, %                          | First-Time Mothers | Male Partners |
|                                 | Baseline (n=338) | Endline (n=339) | Baseline (n=224) | Endline (n=224) |
| 15–19 years                     | 28.7             | 28.3<sup>a</sup> | 1.3              | 0.9              |
| 20–24 years                     | 62.7             | 67.8<sup>a</sup> | 28.6             | 30.4             |
| 25–29 years                     | 1.2              | 2.7<sup>a</sup>  | 40.2             | 38.8             |
| 30 years plus                   | 0.0              | 0.3<sup>a</sup>  | 29.9             | 29.9             |
| Don’t know/missing              | 7.4              | 0.9<sup>a</sup>  | 0.0              | 0.0              |
| Mean age, years                 | 20.6             | 21.1<sup>b</sup> | 27.5             | 27.4             |
| Local government area, %        |                  |                 |                  |                  |
| Ikom                            | 44.4             | 44.0            | 50.9             | 50.9             |
| Obubra                          | 55.6             | 56.0            | 49.1             | 49.1             |
| Marital status, %               |                  |                 |                  |                  |
| Never married                   | 62.7             | 53.1<sup>b</sup> | 30.4             | 31.7             |
| Living with partner/married     | 37.3             | 45.4<sup>b</sup> | 69.6             | 68.3             |
| Divorced/separated/widowed      | 0.0              | 1.5<sup>b</sup>  | 0.0              | 0.0              |
| No. of living children, %       |                  |                 |                  |                  |
| 0                               | 14.5             | 7.7<sup>b</sup>  | 9.8              | 4.5              |
| 1                               | 85.5             | 92.0<sup>b</sup> | 85.3             | 90.2             |
| 2                               | 0.0              | 0.3<sup>b</sup>  | 4.9              | 5.4              |
| Age of youngest child (among participants with at least 1 child) | | | |
| Mean age of youngest child (months) | 6.9 months (n=289) | 8.7<sup>a</sup> (n=312) | 8.8 months (n=199) | 10.9 months (n=214) |
| Residential arrangement, %      |                  |                 |                  |                  |
| Currently lives with partner    | 45.0             | 43.4            | 75.4             | 72.8             |
| Education level, %              |                  |                 |                  |                  |
| Primary                         | 13.9             | 10.9            | 3.6              | 4.0              |
| Junior Secondary (completed)    | 35.2             | 36.6            | 9.8              | 10.3             |
| Secondary (completed)           | 47.6             | 45.4            | 67.9             | 67.0             |
| Polytechnic                     | 1.8              | 2.9             | 4.0              | 3.1              |
| University                      | 1.5              | 4.1             | 14.7             | 15.6             |
| Works to earn money, %          |                  |                 |                  |                  |
| Yes                             | 36.1             | 56.9<sup>a</sup> | 84.8             | 86.6             |

<sup>a</sup> Chi-square P < .000.<sup>b</sup> Chi-square P < .05.
2 months elapsed between data collection efforts for male partners.) In addition, by endline, FTMs were more likely to be in union or married and were more likely to be working to earn money than at baseline. As expected, no significant differences were noted among male partners from baseline to endline with respect to these key demographic variables.

**Birth Spacing Intentions**

One of the key messages of the FTP interventions was to encourage a spacing gap of 3 years or more between births. Figure 1 shows that at baseline, only 17% of FTMs and 40% of male partners at baseline preferred no more children or wished to wait 3 years or longer to have another child. At endline, 81% of FTMs (P<.000, Pearson chi-square) and 88% of male partners preferred no more children or to wait 3 years or longer (P<.000, McNemar’s test). Importantly, an alignment in birth spacing intentions generally occurred for both FTMs and male partners.

**Awareness of Modern Contraceptive Methods**

The FTM and male partner interventions emphasized knowledge and use of postpartum contraception. Knowledge of modern contraceptive methods increased over the course of the interventions, with the percentage of FTMs and male partners who could spontaneously recall at least 3 modern methods increasing significantly (see Figure 2). The percentage of FTMs and male partners who could spontaneously recall at least 3 modern methods nearly doubled over the life of the interventions, increasing significantly among FTMs from 50% at baseline to 94% at endline (P<.000, Pearson chi-square), and among male partners from 38% at baseline to 75% at endline (P<.000, McNemar’s test).

**Myths and Misperceptions of Using Modern Contraception**

A key finding from formative research conducted prior to the FTP interventions was that many FTMs and their male partners believed that using contraception can damage a woman’s reproductive organs and create difficulties in conceiving or can even cause permanent sterility after discontinuation. Thus, most believed that it is best for a woman to use contraception for limiting fertility only after achieving one’s desired family size. Correcting this misconception was an area of focus throughout the interventions. Figure 3 presents the percentage of interviewed participants who held this belief at baseline and endline. At baseline, 55% of FTMs and 29% of male partners agreed that using contraception could negatively affect a woman’s ability to have children in the future. At endline, only 1% of FTMs (P<.000, Pearson chi-square) and 7% of male partners held this belief (P<.000, McNemar’s test).

**Spousal/Partner Approval for Using Modern Contraception**

Husband’s or partner’s approval for using a method of contraception (or perceived approval by FTMs) may be a critical factor in facilitating an FTM who is married or in union to accept and use contraception. Figure 4 presents the percentage of interviewed participants who held this belief at baseline and endline. At baseline, 42% of FTMs and 67% of male partners agreed that their partners supported using contraception. At endline, 82% of FTMs (P<.000, McNemar’s test) and 95% of male partners (P<.000, McNemar’s test) held this belief.

---

**FIGURE 1.** Percentage of First-Time Parents Who Do Not Want Another Child or Who Wish to Wait 3 Years or Longer to Have Their Next Child, Cross River State, Nigeria
contraception. FTMs were asked if they thought that their husband/partner would approve if they wanted to use a method of contraception to space their next child, and male partners were asked if they themselves would approve of their wife/partner using a method of family planning, as well as whether or not they would give her money to seek family planning services. At baseline, only about two-thirds (67%) of the FTM participants thought that their partner would approve of their use of family planning to space their next child, which increased significantly to 80% \((P<.000, \text{Pearson chi-square test})\) at endline (see Table 3). Male partners, however, were much more likely
to approve at baseline (91%), and this did not change significantly at endline (94%). Nearly all (94%) of male partners agreed that they would be willing to support their female partner/wife with money to seek family planning services at endline, which significantly increased from baseline (88%, \( P < .05 \), McNemar’s test).

### Couple Communication on Family Planning

Having discussions with one’s partner or other influential people is often associated with interest in and voluntary use of family planning. The FTP interventions included activities and discussion around partner communication on family planning and birth spacing. Figure 4 presents baseline and endline data on discussions about family planning with partners and other influencers among both FTMs and male partners. Reported discussions about family planning among FTMs (regardless of marital/union status) doubled from baseline (41%) to endline (80%, \( P < .000 \), Pearson chi square) and increased significantly among male partners from 69% to 91% \( (P < .000, \) McNemar’s test). Discussions among FTMs and male partners with other influential people also increased from baseline to endline (data not shown), from 28% to 55% for FTMs \( (P < .000) \) and from 17% to 42% for male partners \( (P < .000, \) McNemar’s test). When asked with whom they discussed family planning in the past 3 months, FTMs were most likely to report discussing family planning with a mother (43%), sister (34%), or friend (51%) at endline \( (n=187, \) data not shown); male partners were most likely to discuss family

### TABLE 3. Percentage Distribution of Partner Support for Family Planning by Participant Group and Baseline/Endline

<table>
<thead>
<tr>
<th>Variable</th>
<th>First-Time Mothers</th>
<th>Male Partners</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline, % (n=338)</td>
<td>Endline, % (n=339)</td>
</tr>
<tr>
<td>Agrees that husband/partner would approve of using family planning to space next child</td>
<td>66.9</td>
<td>79.6 a</td>
</tr>
<tr>
<td>Would give wife/partner money to seek services if she wanted to use family planning to space her next birth</td>
<td>87.9</td>
<td>93.8 b</td>
</tr>
</tbody>
</table>

a Chi-square \( P < .000 \).
b McNemar’s test \( P < .05 \).

**Having discussions with one’s partner or other influential people is often associated with interest in and voluntary use of family planning.**
planning with a friend (73%) or a brother (21%) at endline (n=95, data not shown).

**Couple Decision Making About Family Planning**

Both FTMs and male partners were asked at baseline and endline about decision-making responsibility for using family planning (data not shown). The percentage of FTMs and male partners who reported that they should decide together to use family planning was high even at baseline: 82% (N=338) of FTMs and 88% (n=224) of male partners agreed that using family planning should be a joint decision before the intervention. However, this percentage significantly increased for both participant groups by endline; by the end of the intervention, 96% (N=339; \( P < .000 \), Pearson chi-square) of FTMs and 99% (n=224; \( P < .000 \), McNemar’s test) of male partners agreed that using family planning should be a joint decision. Perhaps even more important, relatively few FTMs reported that husbands/partners were the primary decision maker about family planning, suggesting that contraceptive use was largely voluntary for these young women.

**Current Voluntary Use of Modern Contraception**

The key objective of the FTP interventions was to increase current voluntary use of a modern contraceptive method. Figure 5 shows that current use of a contraceptive method among both FTMs and male partners significantly increased from baseline to endline. Current use increased from 26% (n=288) to 79% (n=316) among nonpregnant FTMs (\( P < .000 \), Pearson chi-square), and from 44% to 81% (n=200) among male partners (\( P < .000 \), McNemar’s test). Importantly, other positive changes in family planning knowledge, attitudes, communication, and decision making all support the overall increase in informed, voluntary contraceptive use by FTP participants.

A logistic regression analysis was also performed to confirm the bivariate findings above, predicting current use of any modern family planning method (implants, intrauterine devices, injectables, oral pills, male or female condoms, emergency contraception, or standard days method) among both FTMs and male partners (in separate models, data not shown). All relevant demographic variables were included in the model, as well as attitudes toward family planning, couples’ discussions about family planning, perceived partner approval and joint decision making for family planning, as well as a variable representing the survey wave (baseline/endline, with baseline as the reference category). This analysis revealed that for both FTMs and male partners, survey wave was highly significant (\( P < .001 \)) with adjusted odds ratios of 3.3 for FTMs and 3.7 for male partners. This means that modern contraceptive uptake significantly increased from baseline to endline for both groups of participants. In other words, after controlling for hypothesized predictors of family planning use, including demographic factors (age, marital status, education level, age of youngest child) and all attitudes related to family planning use presented in this report (including perceived safety of contraceptive methods, partner approval, and decision making related to

**FIGURE 5.** Current Use of a Modern Contraceptive Method (Among First-Time Parents Not Currently Pregnant), Cross River State, Nigeria

![Bar chart showing current use of modern contraceptives among FTMs and male partners](chart.png)
family planning use), FTMs were approximately 3 times more likely and male partners nearly 4 times more likely to be using a modern family planning method at endline, compared with baseline.

**Method of Contraception Used**

Figure 6 shows the type of contraceptive method used among FTM respondents not pregnant at the time of data collection (multiple responses were possible). The graph reveals that use of implants and injectables increased significantly from baseline to endline for nonpregnant FTMs, with implants being the most commonly used method among all respondents (men’s reported use of implants also increased significantly from baseline to endline, but since nearly all male partners had an FTM partner in the program, this information is presented for FTMs only). At baseline, only 17% of FTMs (n=287) reported using an implant, whereas 65% of FTMs (n=316) were using implants at endline ($P<.000$, Pearson chi-square test). Importantly, use of implants aligns with the overall spacing intentions (majority reported 3 or more years) indicated by both FTMs and male partners at endline.

**DISCUSSION**

This article has highlighted selected results of a program designed to address some of the critical barriers faced by young FTMs and their male partners in using family planning to space their second and subsequent children. Implemented by local organizations and resource persons, the interventions aimed to increase HTSP awareness and intentions, build awareness of modern contraceptive methods, dispel key myths and misperceptions about family planning, address gender norms and barriers, increase social and partner support for family planning use, and provide referrals and facility linkages for obtaining specific contraceptive methods. We examined key HTSP and family planning indicators among participating FTMs and their male partners to determine if they successfully changed attitudes and behaviors related to birth spacing and voluntary contraceptive use over the course of the program.

This intervention evaluation included a coordinated baseline and endline questionnaire among a scientific and robust sample of FTM participants and a census of male participants using a trained team of interviewers and digital mobile data collection tools. The results show that the interventions attracted and retained a diverse range of FTMs and male partners (in terms of key sociodemographic variables such as marital status and education) and was successful in improving birth spacing intentions and current use of contraceptive methods from baseline to endline, even after controlling for key sociodemographic and attitudinal variables. Important indicators of contraceptive awareness, attitudes, and couples’ communication increased significantly from baseline.

* $P<.000$.

** $P<.01$.

The interventions were successful in improving birth spacing intentions and current use of contraceptive methods from baseline to endline.
to endline, and also significantly predicted current use of a modern contraceptive method over the course of the intervention. Along with overall increased contraceptive uptake, FTMs and their partners chose to use more effective and/or long-acting methods (injectables and implants), perhaps reflecting, in part, their longer birth spacing preferences, interest in a newly introduced method (implant), as well as the recent increased availability of these contraceptives locally.

These positive findings should be interpreted with several limitations in mind. This study relied on self-reported information gathered during face-to-face interviews, subject to both courtesy and recall bias. This bias was minimized through training of interviewers and design of the questionnaire. Another limitation is related to the possibility that other family planning activities took place in CRS, Nigeria, in the same or nearby geographic areas, which also may have resulted in increased health knowledge and behaviors including a high level of family planning uptake by program participants; however, Pathfinder International’s CRS program staff reported that no other family planning-related partner activities had taken place concurrently in Ikom or Obubra LGAs. In addition, participants might have changed between the pretest and the posttest regardless of the interventions because they are maturing and learning, especially as parents of young infants. This limitation was minimized by ensuring that baseline and endline data collection coincided tightly with program implementation. A final limitation is related to the self-selection of program participants and propensity of more empowered individuals (as opposed to those in the general population) to participate. While it is highly likely that participants had some propensity toward the information and messages received during the program, baseline attitudes and knowledge were consistent with findings from the formative research (described previously).

While previous E2A FTP projects primarily focused on activities with FTMs related to HTSP/family planning, the interventions in CRS engaged male partners more systematically. We included group activities with a subset of nominated male partners of HTM peer group members to address gender norms, increase male engagement in HTSP and family planning, and promote couple communication and joint decision making. Specific activities were included to generate evidence on both the implementation experience and on health and family planning outcomes for FTMs and their male partners emerging from this programming effort. Importantly, endline results show that FTMs and their partners were generally aligned on key family planning attitudes and birth spacing intentions, which may have facilitated increased contraceptive use and method choice. These results suggest that couple-oriented interventions or joint activities can work well—even in a context in which many FTPs are not in formal unions or necessarily living in the same household.

Our experience in CRS suggests that FTMs and male partners may be particularly open to HTSP and family planning use because they face the practical and financial realities of raising a child. Community-based resources like CBOs, CVs (or similar community health workers), peer leaders, and others provided FTPs with tailored and timely access to information and services, as well as linkages with health facilities critical to ensuring access to a full range of needed services. Many FTMs do not routinely access health facilities or may not be ready to consider family planning options when they do. Therefore, such approaches may work better than only integrating family planning into clinic-based services (e.g., postpartum family planning, postabortion care services, or even family planning integration into antenatal care), especially where there are inequities in access to and use of health care by young FTMs.

While our project included multiple interventions focused on FTPs, all activities were implemented over a 4-month period through existing health facilities and a local CBO, using trained and certified (but not yet employed) CHEWs. Several elements, such as home visits, community outreach, and the provision of modern contraceptive methods, were already included within the general mandate of the primary health care system. Building on existing community- and facility-based resources to identify and reach FTPs with tailored activities generated compelling results and provided a model that can be adapted based on available resources and scaled-up across the state.

**CONCLUSION**

The E2A experience in CRS shows that tailored interventions with FTPs can achieve important HTSP and family planning results within a relatively short time frame. FTMs and partners are coping with multiple challenges as new parents and are receptive to information and options that allow them to delay subsequent births. The emerging high demand for family planning across diverse
FTMs and partners—especially for more effective and longer-acting contraceptive methods—underscore the importance of engaging FTPs during this critical moment in their reproductive lives. In particular, the CRS experience suggests 3 essential program elements: (1) ensuring the availability of modern contraceptive methods (especially implants) through local health facilities; (2) using locally based resource persons or community-based health workers to conduct home visits with FTPs to provide tailored health information and referrals, as well as build linkages with the formal health sector; and (3) using activities that address gender norms and couple dynamics to foster better alignment, communication, and joint action on reproductive issues. All activities can be implemented through locally based resource persons, who are best positioned to identify and reach young FTPs of different characteristics and situations. The results that can be achieved, along with high levels of engagement from FTPs, demonstrate the importance of investing in these types of interventions, ideally addressing all priorities for family planning, reproductive health, and maternal, neonatal, and child health across the FTP lifespan, from pregnancy through the postpartum period.

Acknowledgments: The authors wish to thank the Cross River State Government, Nigeria, for its support of this work and the project team from Pathfinder International, Nigeria, including Farouk Jega, Yemisi Femi-Plus, Nenka Alabi, and Benjamin Asemota. We also thank the Greater Hands Foundation executive director, Arit Williams, and her team. Researchers from Research and Communications Services Limited conducted baseline and endline surveys with project participants, under the direction of Akinsewa Akiode, principal consultant. We also acknowledge Connie Lee, Laurel Lundstrom, Eric Ramirez-Ferrero, Caitlin Thistle, and Cary Wornell for their critical and valuable review of this manuscript and Marylene Wumakoya and Maurice Mutsiya for their helpful advice on statistical analysis. Finally, and most importantly, we thank the hundreds of young first-time mothers and their partners in Cross River State.

Funding: This publication was made possible through support provided by the Office of Population and Reproductive Health, Bureau for Global Health, U.S. Agency for International Development, under the terms of Award No. AID-OAA-A-11-00024. The opinions expressed therein are those of the author(s) and do not necessarily reflect the views of the U.S. Agency for International Development.

Competing interests: None declared.

REFERENCES
19. Williamson LM, Parkes A, Wright D, Petticrew M, Hart GJ. Limits to modern contraceptive use among young women in developing
Peer Reviewed

Received: March 13, 2020; Accepted: October 6, 2020; First published online: November 19, 2020


© Morgan 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-20-00111
Routine Family Planning Data in the Low- and Middle-Income Country Context: A Synthesis of Findings From 17 Small Research Grants

Bridgit Adamou, Janine Barden-O’Fallon, Katie Williams, Amani Selim

Key Findings

- We found 4 major themes affecting family planning data quality, analysis, and use:
  - The enabling environment for managing and using family planning information
  - Barriers to integrating family planning in routine health information systems
  - Gaps in analyzing, interpreting, and using routine family planning data
  - Family planning data use in decision making

- Systematic, organizational, cultural, and technical barriers affect data quality and limit subsequent analysis, interpretation, and use of information.

Key Implications

- Program implementers should consider:
  - Standardizing family planning indicators across sectors (public and private) and data collection tools
  - Conducting regular staff trainings and capacity building to improve data literacy, collection, and reporting
  - Investing in the human and technological resources needed for effective data collection, analysis, and use.

- Policy makers should:
  - Enact and commit to continuous financial support
  - Emphasize well-defined data collection and reporting processes, including clearly defined indicators and harmonized data collection tools
  - Provide well-supported technical infrastructure.

ABSTRACT

Health information systems rely on high-quality data to measure, track, and inform decision making. Currently, the quality, uptake, and use of family planning data in routine health information systems is limited, presenting an opportunity for improvement on many levels. The current synthesis assessed findings from 17 small grants that MEASURE Evaluation issued to low- and middle-income country research teams between 2015 and 2019. Main findings from that research were collaboratively categorized in 4 major themes: (1) the enabling environment for managing and using family planning information; (2) barriers to integration of family planning in routine health information systems; (3) gaps in the analysis, interpretation, and use of routine family planning data; and (4) family planning data use in management, programmatic, and budgetary decisions. Data quality at the systemic, organizational, technical, and output levels was a crosscutting theme. Collectively, the findings outline barriers to and opportunities for improved integration of family planning data and subsequent strengthening of routine health information systems.

BACKGROUND

The provision of health care services and information about their quality and quantity are critical components of a health system. These components must function together to strengthen service delivery programs and improve population health. Countries use health information systems (HIS) to measure and track health services, allowing them to plan, evaluate, and implement health strategies. An efficient HIS draws from multiple levels of the health system, using clearly defined indicators, up-to-date standards and guidelines, accessible data collection and analysis tools, and stakeholder collaboration and support to enable evidence-informed decision making. A key component of an HIS is a routine health information system (RHIS), fundamentally composed of indicators to track management information needs and data collection, transmission, processing, and analysis, which should all lead to information use. Data from RHISs include service statistics, management and logistics data, and financial data, and provide information on client health status, facility and budgetary capacity, and services and resources.
administered or available.4 These RHIS data constitute the main pillar for monitoring service delivery programs at the national level in low- and middle-income countries (LMICs).5 Despite a sound framework for an effective HIS, earlier research found underperforming RHISs due to several factors, such as poor data quality; indicators lacking standardization, clear definitions, and accurate calculations; inadequate electronic data capture and reporting; incomplete data analysis; poor management support; and weak use of information for planning and decision making.4 A strong RHIS that supports data-informed decisions requires 4 key actions: regularly assessing the organizational, technical, and behavioral factors that affect decision making to improve data demand and use; engaging data producers (those who design and manage research and information systems) and data users (those who use data in program improvement and development) in the decision-making cycle; improving data quality; and improving data availability, defined as data synthesis, data communication, and access to data.5

For many LMICs, accurate collection, reporting, analysis, and use of routine data from an HIS are challenging tasks that span health areas, from maternal and child health to infectious and chronic diseases.7 It is also a challenge for LMICs to ensure that routine family planning data in their HISs are accurate and complete. The family planning community has paid relatively little attention to strengthening RHISs, causing the field to fall behind other health areas.8 Recent efforts to collect data for the FP2020 global initiative have brought increased attention to family planning service statistics, data quality, and reporting mechanisms.8

Despite the recent attention focused on family planning in RHISs, the production of high-quality information sufficient for program planning, monitoring, advocacy, and other decision-making needs has proven difficult. Health care providers that do collect routine family planning data often find that the larger HIS into which these data feed lacks the appropriate reporting or synthesis mechanisms; in other cases, the family planning data are of poor quality or are not collected consistently.5 Knowledge gaps related to routine family planning data include how to improve the quality of family planning data, how to address barriers to integrating family planning data in RHIS, and how to encourage analysis and use of the data to improve family planning outcomes.

To better understand the dynamics of family planning data collection, integration, and use, the MEASURE Evaluation project, funded by the U.S. Agency for International Development (USAID), provided technical and financial support for researchers in LMICs to investigate issues related to the collection, aggregation, and use of routine family planning data. This article synthesizes the family planning-specific research results from 17 small grant-funded projects, organized by common themes, to shed light on the status of family planning in RHISs.

## METHODS
In 2014, MEASURE Evaluation implemented a program funded by the USAID Office of Population and Reproductive Health that provided small grants for research related to the collection, analysis, and use of routine family planning data in 24 priority countries. The overarching goal of the program was to produce evidence that could help improve RHISs and advance family planning outcomes. The MEASURE Evaluation small grants program aimed to (1) address research gaps in routine health information for family planning/reproductive health (RH) to inform policy and programmatic decision making, (2) strengthen research capacity among local agencies, and (3) increase use of research findings by providing an opportunity for the data to be disseminated to and used by local stakeholders to inform decision making. The program supported both primary and secondary data collection and analysis. Grant recipients were required to secure appropriate ethical review and approval prior to research implementation. Five rounds of awards were implemented over a 5-year period (2015–2019), generating 360 applications and resulting in 19 funded research projects in 11 countries (Table). Recipients represented a mix of university, quasi-governmental, nonprofit, and private research organizations. The grant amounts ranged from US $10,000 to US$24,000 in direct funds, with an average award of US$14,400. We required recipients to complete a technical working paper of their research results and to conduct at least 1 data use activity with stakeholders (such as the presentation of findings at technical working group meetings, workshops, or conferences). We provided technical assistance as needed throughout the application, implementation, writing, and dissemination stages of the research projects. Details about the program were previously described by Adamou.10

To synthesize the results of these research projects, we reviewed the 19 small grants working papers, excluding 2 from the synthesis because
<table>
<thead>
<tr>
<th>Research Organization</th>
<th>Research Title</th>
<th>Study Objective(s)</th>
<th>Geographic Coverage</th>
<th>Data Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrated Health Initiative</td>
<td>Integrating Family Planning Data from Public and Private Health Facilities in Malawi: How Current Approaches Align with FP2020 Goals</td>
<td>Find approaches to improve the national Health Information System by integrating family planning data from private-sector service delivery points and government facilities</td>
<td>2 districts in each of the 3 regions in Malawi</td>
<td>Desk reviews of all national policy documents guiding family planning data and data collection; field observations; 71 KIs with staff from national-level institutions of the MOH, zonal offices of the 5 quality control divisions (i.e., zones) in Malawi, and family planning service providers, HMIS officers, health surveillance assistants, family planning coordinators, and data clerks</td>
</tr>
<tr>
<td>Rivers State of Nigeria Primary Health Care Management Board</td>
<td>Use of Technology to Manage Health Data in Rivers State, Nigeria: A Qualitative Study on Family Planning and Routine Health Information Systems</td>
<td>Explore the experiences and perceptions of family planning providers and health information officers on implementing technology for district health data collection and identify factors that affect the sustainability of using technology for data management in Rivers State, Nigeria</td>
<td>Rivers State, Nigeria</td>
<td>21 IDIs with state- and LGA-level HMIS officers, desk officers, monitoring and evaluation officers, and reproductive health coordinators; 2 FGDs with 35 facility health information officers and family planning providers</td>
</tr>
<tr>
<td>Africa Field Epidemiologic Network</td>
<td>Family Planning Indicators Assessment and Data Quality Audit in Selected Health Facilities Across Nigeria</td>
<td>Estimate family planning indicator performance at the health facility level from the HMIS not reflected in DHIS2 to determine the quality of family planning data at the facility level and identify challenges to family planning program implementation in sampled health facilities in Nigeria</td>
<td>2 LGAs in each of the following 6 states in Nigeria: Bauchi, Delta, Enugu, Kano, Osun, and Nasarawa</td>
<td>Administration of a questionnaire via interviews with 114 family planning/reproductive health focal people in selected facilities; 42 KIs with family planning stakeholders and key decision makers in the family planning/reproductive health units at the LGA and state levels in the selected states; 6 FGDs with health workers/service providers</td>
</tr>
<tr>
<td>The Rescue Initiative-South Sudan</td>
<td>Analyzing, Interpreting, and Communicating Routine Family Planning Data in South Sudan</td>
<td>Explore how effectively family planning data in the RHIS are analyzed, interpreted, and communicated, and discuss barriers to RHIS data use and ownership in 2 states in South Sudan</td>
<td>17 counties in 2 states in South Sudan: Central Equatoria and Western Equatoria</td>
<td>Direct observation at service delivery points, individual questionnaires administered to health facility staff, and KIs with a total of 180 study participants</td>
</tr>
<tr>
<td>University of the Punjab, Institute of Social and Cultural Studies</td>
<td>The Routine Health Information Systems in Punjab Province, Pakistan: Exploring the Potential for Integrating Health Information Systems for Family Planning Data</td>
<td>Review the RHIS in Punjab province of Pakistan and explore the potential for integrating community-level data into the national HMIS, particularly family planning data, collected by public or private, for-profit, and not-for-profit organizations</td>
<td>Punjab province, Pakistan</td>
<td>Document review and 16 KIs with lady health workers, the Population Welfare Dept., Rahnuma-Family Planning Association of Pakistan, DHIS office, United Nations Population Fund, and United Nations Children’s Fund</td>
</tr>
<tr>
<td>Research Organization</td>
<td>Research Title</td>
<td>Study Objective(s)</td>
<td>Geographic Coverage</td>
<td>Data Sources</td>
</tr>
<tr>
<td>-----------------------</td>
<td>----------------</td>
<td>--------------------</td>
<td>---------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Department of Population Studies, Makerere University</td>
<td>Integrating Family Planning Data in Uganda’s Health Management Information System</td>
<td>Investigate the facilitators, best practices, and barriers of integrating family planning data into the district and national HMIS in Uganda</td>
<td>Kampala, Jinja, and Hoima districts, Uganda</td>
<td>16 KIIs with MOH officers, HMIS focal persons at non-governmental organizations, HMIS focal persons who were district biostatisticians or medical records officers, and providers who were medical records officers at public and private health facilities; a multi-stakeholder dialogue workshop comprised of 11 participants; and a systematic review of the HMIS in sub-Saharan African countries that are United States Agency for International Development family planning priorities</td>
</tr>
<tr>
<td>International Centre for Diarrhoeal Disease Research, Bangladesh</td>
<td>Using DHIS 2 Software to Collect Health Data in Bangladesh</td>
<td>Explore the perceptions and experiences with using DHIS2 to collect and analyze reproductive, newborn, maternal, and child health data in Bangladesh and to identify facilitators and barriers to using these data at different levels of the health care system</td>
<td>Khulna and Chittagong districts in Bangladesh</td>
<td>Document review; 23 IDIs with community health care providers, nurses, health inspectors, and upazila statisticians; 2 FGDs with district statisticians; and 11 KIs with health managers, HMIS experts, and key decision makers</td>
</tr>
<tr>
<td>Research and Development Division, Ghana Health Service</td>
<td>Experiences and Perceptions of Health Staff on Applying Information Technology for District Health Data Management in Ghana</td>
<td>Explore and document the experiences and perspectives of health staff and managers in the 4 districts on use of mobile technology to collect and manage health data in district health systems</td>
<td>4 administrative districts in Ghana’s Central Region</td>
<td>KIs with 160 frontline health staff (midwives, community health nurses, health information officers, general nurses, and physician assistants) at both the district and subdistrict levels and 14 district and regional health managers and policy makers</td>
</tr>
<tr>
<td>Centre of Population, Health and Nutrition Services</td>
<td>Improving Family Planning Service Delivery in Ghana</td>
<td>Map out the distribution of all family planning service providers in the region and document how the community-based family planning information system is linked to the national system to recommend strategies for supporting program planning and implementation and improving family planning services</td>
<td>Upper East Region, Ghana</td>
<td>Records review and data extraction from DHIS2; survey of all types of service providers in the region’s 13 districts by interviewing the family planning providers present (435) using a structured interview questionnaire; 2 FGDs with the district health management team, staff from different subdistrict health teams, and community health officers</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Research Organization</th>
<th>Research Title</th>
<th>Study Objective(s)</th>
<th>Geographic Coverage</th>
<th>Data Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Governance Links Tanzania</td>
<td>Strengthening Tanzania’s Routine Health Information System: Incorporating Family Planning Quality Assessment Indicators²¹</td>
<td>Investigating the benefit of incorporating indicators related to family planning quality assessment in a decentralized RHIS in rural farming districts around Lake Victoria</td>
<td>Administrative district of Magu, Mwanza Region, in the Lake Victoria zone of Tanzania</td>
<td>Literature review; questionnaire-facilitated individual interviews with 50 health service providers and community health workers; 12 KIs with health service providers, pharmacy staff, civil society organization staff, council health management team members, and district health information officers; 2 FGDs with 40 health service providers and community health workers</td>
</tr>
<tr>
<td>Matibabu Foundation</td>
<td>Integrating Family Planning Data in Kenya’s DHIS ²²</td>
<td>Investigate integration of family planning data in DHIS2, the factors related to lack of integration, and ways to remedy the lack of integration</td>
<td>Siaya and Nairobi counties, with a pretest conducted in Kisumu county in Kenya</td>
<td>Eight KIs with MOH officers from Siaya and Kisumu counties and a representative from the Division of Health Information Systems, at the national level. Four FGDs were conducted with clinicians, nurses, health records officers, and information officers from both public and private health facilities at all levels, from the primary level to county referral hospitals.</td>
</tr>
<tr>
<td>Equitable Health Access Initiative</td>
<td>The Strongest Motivators for Using Routine Health Information in Family Planning: A Prospective Study in Lagos, Nigeria²³</td>
<td>Bridge the knowledge gap concerning the motivators behind using routine health information in family planning to improve the use of family planning services</td>
<td>3 LGAs of Lagos state, Nigeria</td>
<td>12 KIs and 425 questionnaires with men and women working in the health sector</td>
</tr>
<tr>
<td>Afya Research Africa</td>
<td>Family Planning Services in Kenya During a Transition: Utilization Trends Across Counties²⁴</td>
<td>Estimate the general prevalence of family planning use among women of childbearing age and the prevalence of family planning use by county; analyze the trends in family planning utilization over the period of transition, from 2012 to 2015; and estimate the extent to which counties had integrated reporting of family planning services in Kenya’s DHIS²</td>
<td>Kenya</td>
<td>National family planning–related DHIS² data and Kenya Demographic and Health Survey 2014 data</td>
</tr>
</tbody>
</table>
the research topics were not specifically related to an RHIS. The main findings of the 17 remaining papers were extracted, reviewed, and organized by key concepts through an iterative process in which all co-authors participated. Themes were developed around the key concepts. Once organized, the findings within each theme were compared and contrasted. We then summarized the results to present main findings for each theme and to contribute to an overall understanding of current strengths, issues, and gaps in family planning data and RHISs in LMICs.

**TABLE. Continued**

<table>
<thead>
<tr>
<th>Research Organization</th>
<th>Research Title</th>
<th>Study Objective(s)</th>
<th>Geographic Coverage</th>
<th>Data Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mzumbe University, School of Public Administration and Management</td>
<td>Creating a Culture of Data Use in Tanzania: Assessing Health Providers’ Capacity to Analyze and Use Family Planning Data&lt;sup&gt;25&lt;/sup&gt;</td>
<td>Understand health providers’ capacity to analyze collected family planning data and to document available evidence of health service providers using the collected data in their planning processes</td>
<td>2 LGAs within each of the following regions in Tanzania: Lindia, Geita, and Arusha</td>
<td>13 IDIs with facility in-charges, reproductive and child health in-charges, data clerks, and family planning facility-based providers; 2 FGDs with 24 health providers; and non-participant observation in 12 health facilities</td>
</tr>
<tr>
<td>Health Promotion Tanzania</td>
<td>Enhancing Use of Routine Health Information for Family Planning to Influence Decision Making in Tanzania&lt;sup&gt;26&lt;/sup&gt;</td>
<td>Explore the type of family planning information collected, how the data are analyzed, and how the information informs planning and budgeting. It examined ways data are handled across all 5 levels of the health system (i.e., national, regional, district, ward, and village) and when and how the data are utilized.</td>
<td>Kilimanjaro and Mara regions of Tanzania</td>
<td>31 KIIs with health officers in charge, points of contact for family planning or reproductive health and child health, district medical officers, health governance committee, HMIS focal people, and health secretaries from a regional hospital, district hospital, health center (ward level), and dispensary (village level)</td>
</tr>
<tr>
<td>Association for Reproductive and Family Health</td>
<td>Use of Routine Health Information to Inform Budgetary Allocation for Reproductive Health in Cross River State, Nigeria&lt;sup&gt;27&lt;/sup&gt;</td>
<td>To understand the budget process within the state MOH and in the health department of the Calabar municipal local government council; examine the use of routine health information as evidence for budgetary allocation for reproductive health and family planning; identify barriers and constraints to routine data use; explore possible solutions; and dialogue with the stakeholders on how routine health data can be used in the budget process</td>
<td>Calabar Municipal LGA in Cross River state, Nigeria</td>
<td>Desk review of existing family planning data in Cross River State and Calabar Municipal LGA, KIIs with staff from relevant ministries, and questionnaires administered to middle- and junior-level officers at the state and LGA levels</td>
</tr>
<tr>
<td>Access Global Ltd.</td>
<td>Uganda’s Resources to Finance Family Planning Commodities: Implications for a Total Market Approach&lt;sup&gt;29&lt;/sup&gt;</td>
<td>Understand the extent to which in-country resources can mitigate financing shortages for family planning commodities in Uganda, and the implications of a total market approach</td>
<td>Uganda</td>
<td>Literature review; retail audits in 16 pharmacies in Mukono district; and 6 researcher-administered questionnaires with family planning program managers</td>
</tr>
</tbody>
</table>

Abbreviations: DHIS, District Health Information Software; FGD, focus group discussion; HMIS, health management information system; IDI, in-depth interview; KII, key informant interview; LGA, local government area; MOH, ministry of health; RHIS, routine health information system.
RESULTS

The synthesis of results yielded the following main themes: (1) the enabling environment for managing and using family planning information; (2) barriers to integration of family planning in RHISs; (3) gaps in the analysis, interpretation, and use of routine family planning data; and (4) use of family planning data in management, programmatic, and budgetary decisions. All papers discussed the issue of data quality—the systematic, organizational, cultural, and technical barriers that contributed to data quality problems and the effects of poor data quality on analysis, interpretation, and use of information. For this reason, data quality was considered to be a crosscutting theme, and we incorporated it, as appropriate, in each of the 4 thematic areas.

Theme 1: The Enabling Environment for Managing and Using Family Planning Information

The first theme identified in the review of the small grant-funded research papers was related to the enabling environment for the management and use of family planning information. We used the following definition for enabling environment: strong HIS governance and leadership; policy and framework compliance; appropriate resources, such as staffing, technology, and tools; and cross-sector engagement of actors, including private and public entities.11 The small grant–funded reports illustrated how challenges in the enabling environment affected data collection, assessment, and use at all levels.

HIS Governance and Leadership for Compliance

The review indicated that the strength of system governance can be gauged by a country’s ability to enforce its reporting policies and guidelines. Study findings from Malawi, Nigeria, and South Sudan revealed noncompliance and inconsistent submission of family planning data to the national HIS.12–15 Weak governance structures were reflected by countries’ inability to enforce guidelines. For example, despite the protocol in Malawi that private franchises must submit their monthly data summary reports to the district health office, private providers felt no obligation to do so.12 One study participant shared:

When we have compiled the data each month we have a summary, and that summary is sent to our headquarters. Yeah, that’s all, it’s sent to our headquarters. The government has never asked me; of course, I have never sent them any data, no. —Private health service provider

In Pakistan, several private facilities are not legally registered, so it is difficult to collect routine health information from them.16 However, researchers in Uganda found that because the Ugandan Ministry of Health mandates regular submission of HIS reports to health districts as a requirement for private facilities’ renewal of licensure, private and nongovernmental organization health facilities have greater participation in the HIS.17 Furthermore, private nonprofit health facilities (such as faith-based health centers) performed better than public facilities with respect to submission of data because of strict rules enforced by their governing institutions.11

Appropriate Resources

Researchers in Bangladesh identified a shortage of human resources, frequent version changes in the District Health Information Software, version 2 (DHIS2) platform, negative attitudes about electronic data capture systems from some staff, and reliance on donor support as structural barriers to the success of the HIS.18 Consequently, users of the system suggested strong government commitment, deployment of data-quality checks, and accessible technology, along with extensive, sustained financial support, to make the nationwide implementation of the electronic system successful.18

The review also found that a consistent factor in managing an RHIS and the subsequent enabling environment for family planning information was the use of new HIS technologies as an important resource for data capture and reporting. Although the reports mentioned several types of systems, many of the national HIS included a web-based application for electronic data management that was accessible through electronic devices with browser and Internet access. Typically, this application was DHIS2. Research in Uganda found that DHIS2 was considered appropriate and user friendly, and the web-based reporting eased the sharing of health data with stakeholders.17 Researchers in Ghana found that mobile tools enhanced job performance, the quality of data collection, and the efficiency of data management.19 A study participant shared the following:

I can now sit in my office and monitor activities at the peripheries and even at hard-to-reach areas, which activity would otherwise have cost transport, fuel, and much time. Now, I can go on [the mobile technology] and check ... everywhere a health facility is located, or
a health staff may work with ease using technology.

—District-level health manager

Nevertheless, the implementation of new technology hindered progress when necessary resources and infrastructures were inadequate. For example, one-third of the 435 family planning service delivery points surveyed in the Upper East region of Ghana did not have electricity, making electronic data very challenging.20 Research from Rivers State, Nigeria found the new technology led to parallel systems. Health facilities reported family planning data into DHIS2, but system users continued to use paper-based data collection tools at the health facilities because of logistical challenges with the electronic infrastructure including frequent power outages, hardware problems, broken mobile devices, and lack of Internet connectivity.13 Nearly all (96.6%) of the study participants in the Central region of Ghana concurrently used paper-based data collection and reporting tools and mobile technology for collecting and transferring health data.20 The research teams in Bangladesh and Tanzania found similar barriers.18,21 Additionally, the researchers in Rivers State, Nigeria reported faulty computer equipment, inadequate training on use of data tools, and low levels of information and communication technology skills.13 Study participants complained of substandard government-issued mobile devices and difficulties using mobile phones for data collection13:

Some of us are not so perfect with the phone, because, eh, at our local government area, we find it difficult to send the message on the phone. But when you get to where you can connect to the Internet, they say “no service.” You will continue waiting, waiting, waiting until you are fed up. At the end of the day, the phone itself, which we are given to serve at the health facility, remains faulty. So, it wasn’t so adequate with us. —Health information officer at public primary health center

Another example of inadequate resources to support an enabling environment was insufficient funding to support district health offices. This translated into scarce resources needed for a fully functioning HIS, such as data collection guidelines; computers and mobile devices; paper record books and forms; and HIS staff available for data consolidation, verification, analysis, and supportive supervision.17,20,22 A district-level study participant in Ghana said20:

I am one person in this office who enters reports from all these facilities into the system, who does data assessment, who analyzes, validates, and everything.

Cross-Sector Engagement

The often-dissonant relationship between public and private health care sectors played a large role in stratifying data collection and limiting information sharing. Even public and private service providers who operated in the same data catchment space often used separate protocols, separate planning procedures, and data collection mechanisms that were not standardized.9,16,20 The differing approaches to family planning data collection and reporting weakened data sharing in the absence of collaborative networks. Study respondents in Malawi estimated that less than half of the data generated in the private health facilities were reported.12 Although a system existed to flow data from the facility level to the national HIS, major issues with private-sector actors (e.g., noncompliance, inconsistent data submission, poor-quality data, and reporting delays) prevented interpretation of these data.12 The study in Pakistan reflected a similar culture of noncompliance and noncooperation.16 In contrast to these findings, research in Uganda found that collaborative networks existed between donor-funded implementing partners and local organizations, enabling training, financial support, and technical assistance in designing data collection tools essential for better HIS performance and sustainability.17 This was seen as an opportunity to improve public–private facility interaction by strengthening and standardizing reporting requirements.17

Theme 2: Barriers to Integration of Family Planning in RHISs

The second theme that emerged from the review centered on barriers to the integration or inclusion of family planning as a health area in RHISs. Generally, the studies revealed poor data flow from the service delivery points to the district and national HISs; challenges with implementing data collection tools; lack of clear, standardized family planning indicators; and disjunctive networks of collaboration as limitations to the full integration of family planning in RHIS. Many of the studies revealed incomplete integration of family planning data along the designated data-flow chains, and discrepancies existed between mechanisms for data collection and management at the national, community, and facility levels.17,20,22 For example, research in Kenya revealed that the paper-based national data summary tool, known as the MOH 711, which is used as a template to transfer data to DHIS2, includes family planning methods that are not recorded in either family planning...
registers or DHIS2. A health official in Kenya remarked, “I know there is no specific one [tool] for family planning that is really standard for all.”22 This lack of data harmonization creates ambiguities in the system, compromises data quality, and makes the family planning situation incomplete.23,24 Multiple studies found discrepancies in the ability to collect and record family planning data specifically at the facility or community level.15,20,22 In Ghana, there were no required reporting mechanisms for certain community-level family planning service providers, such as pharmacies and licensed chemical sellers.19 Similarly, the HIS in Pakistan does not have a mechanism to record both community- and facility-based family planning services for each client.16 Because the country’s management information systems (the DHIS and commodity logistics management information systems) are managed by different departments, integrating the systems will require high-level organizational restructuring.

As suggested in Theme 1, issues with technical infrastructure, such as mobile and web-reporting challenges, and restricted access to computer-based systems negatively affected data integration and flow.12,17,20 For example, in Kenya, data entry and editing rights are restricted to the subcounty health records and information officers. This restriction hinders service providers’ ability to efficiently and effectively record family planning data, which ultimately affects what is captured in DHIS2.22 A study respondent explained the problem:

The task sometimes overwhelms the staffs, who would end up with forgetfulness. The notion of I’ll tally tomorrow, and again, tomorrow comes—I’ll tally the next day. So, it is continuous. When you come back tallying at the end of the month, you end up tallying wrong information. Your addition might not be right, so you find discrepancies in data. DHIS2 is not the same as data in the facility. This has happened several times. We even have this report last week, during review meeting, and underreporting—to mean what we have on the ground is not what we have at DHIS2. It’s either due to shortage of staffs, or somebody is not able to fill in data at the right time. The ideal is, one should give the service and then tally real time, then give the document by the end of the day tally. —Facility in-charge at public health facility

Organizational factors, such as a failure to prioritize family planning data, also influenced integration into the RHISs. Research from Pakistan reflected this prioritization problem; although an RHIS existed for various health care entities, public departments and nongovernmental organizations did not regularly report family planning data into it.16

Insufficient human resources for both provision of services (and therefore data capture) and supportive supervision and feedback, too few data collection tools (i.e., computers, tablets, forms, and family planning record books), incorrect data entry, and lack of harmonization of data collection tools also affected the inclusion of family planning data into the RHIS. Problems with data collection tools included electronic and paper-based forms without family planning indicators, improper report consolidation, and unavailable collection mechanisms.16,17 Additionally, many health facilities involved in these studies operated both with paper-based patient registers and electronic systems, and these disjointed methods led to missing or incomplete data entry—a problem that was compounded by a lack of training for data collectors and a lack of supportive supervision.17,20,22 For example, when forms are revised, not all family planning providers are trained on the changes, which exacerbates the problem of low data literacy and results in family planning data being excluded from the HIS. A district-level health officer in Uganda revealed17:

I have never heard of nurses and midwives going for refresher training on family planning data in the HMIS [health management information system].

Poor integration of family planning data into the RHIS also stemmed from the limited pool of standardized family planning indicators both in health facility registers and the national HIS.19,23 In Kenya, researchers found that weak indicators at the facility level affected summary data compiled at the intermediary ministerial level, in turn limiting tertiary indicators in the national HIS.22 Without well-defined, standardized indicators harmonized across the HIS, the data collection tools fell short in recording family planning practices and services. The study in Pakistan found that this data shortcoming spurred provider dissatisfaction with the existing family planning indicators.17 Data collection forms did not provide indicator definitions or a place to record changes in family planning choice by individuals.15 Indicator limitations led to such data-quality issues as inaccuracy, overreporting, and missing family planning measurements.15

Theme 3: Gaps in Analysis of Routine Family Planning Data
The third major theme of the review related to gaps in analysis of routine family planning data.
All the studies underscored that problems or the perception of problems with data quality and reliability resulted in limited analysis and use of routine family planning data. For example, Tanzanian researchers found that more than 90% of their study respondents agreed that a big limitation in assessing routine family planning data was poor-quality data (another being the lack of financial resources to support the collection of high-quality routine data). The limited analysis of routine data was also mentioned as a result of a lack of training on electronic data capture tools, a lack of data literacy among system users, poor data analysis skills, overburdened human resources, and an absence of leadership or guidance for family planning data analysis.

The researchers found that there was often an awareness, but not a full understanding, of family planning indicators and their ability to accurately capture intended information, hampering the appropriate analyses. For example, when researchers in Tanzania asked study participants (e.g., family planning service providers, HMIS officers, district medical officers, facility in-charges) to identify the source of family planning indicator data, nearly 20% did not acknowledge men to be a source of family planning information, and one-third did not think any family planning data were obtained from youth.

Many of the studies outlined mechanisms through which family planning data-capture tools might be used to improve data quality and thereby improve data analysis. Researchers in Tanzania recognized that incorporating explicit quality assessment indicators (such as quality of care or attitudes toward family planning) for family planning data into routine data collection could strengthen the usefulness of facility-level data when qualitative and quantitative indicators are analyzed together. The study authors added that an additional pathway for improved data quality and reliability was to explore and invest in technology options for data capture and transmission that were appropriate and cost-effective for rural settings and facilitated easier data analysis. In Nigeria, it was suggested that integrating family planning services in other health areas, such as HIV, immunizations, delivery, and postabortion care, could improve family planning data quality and reliability, and therefore analysis and interpretation, by creating a more complete picture of which family planning services are provided where and to whom.

Findings revealed that many management, programmatic, and budgetary decisions were not informed by evidence. All the research papers underscored that problems, or the perception of problems, with data quality and reliability resulted in limited analysis and use of routine family planning data. For example, Tanzanian researchers found that more than 90% of their study respondents agreed that a big limitation in assessing routine family planning data was poor-quality data (another being the lack of financial resources to support the collection of high-quality routine data). The limited analysis of routine data was also mentioned as a result of a lack of training on electronic data capture tools, a lack of data literacy among system users, poor data analysis skills, overburdened human resources, and an absence of leadership or guidance for family planning data analysis. The researchers found that there was often an awareness, but not a full understanding, of family planning indicators and their ability to accurately capture intended information, hampering the appropriate analyses. For example, when researchers in Tanzania asked study participants (e.g., family planning service providers, HMIS officers, district medical officers, facility in-charges) to identify the source of family planning indicator data, nearly 20% did not acknowledge men to be a source of family planning information, and one-third did not think any family planning data were obtained from youth.

Many of the studies outlined mechanisms through which family planning data-capture tools might be used to improve data quality and thereby improve data analysis. Researchers in Tanzania recognized that incorporating explicit quality assessment indicators (such as quality of care or attitudes toward family planning) for family planning data into routine data collection could strengthen the usefulness of facility-level data when qualitative and quantitative indicators are analyzed together. The study authors added that an additional pathway for improved data quality and reliability was to explore and invest in technology options for data capture and transmission that were appropriate and cost-effective for rural settings and facilitated easier data analysis. In Nigeria, it was suggested that integrating family planning services in other health areas, such as HIV, immunizations, delivery, and postabortion care, could improve family planning data quality and reliability, and therefore analysis and interpretation, by creating a more complete picture of which family planning services are provided where and to whom.

**Theme 4: Family Planning Data Use in Management, Programmatic, and Budgetary Decisions**

The final theme identified in this review was family planning data use in management, programmatic, and budgetary decisions. Despite issues with data quality and reliability, routine family planning data were sometimes key for programmatic decision making. For example, in northern Tanzania, RHIS data were perceived to be an effective and important resource in decision making for improving family planning services. A member of a council health management team said:

> RHIS is a very important tool to us in council health management teams. We depend on it to make important decisions to improve health services in terms of understanding demand and resource allocation.

However, the findings revealed that many management, programmatic, and budgetary decisions were not informed by evidence. For example, researchers in Nigeria found that despite the high unmet need for family planning (30.8%) in Cross River State, only 0.1% of the state’s health budget was earmarked for RH and family planning in 2014. (For comparison, in 2009–2010, RH represented 13.9% of total health expenditures in Kenya.) In one case, the necessary data were not available: in Uganda, the National Medical Stores, development partners, and implementing partners were unable to access data on the quantity of family planning commodities imported and the cost price because the National Drug Authority did not have the data in retrievable form, even though organizations required this information for calculating budgets and funding needs. Use of the data for decision making often did not occur at lower levels of the system either.

Several factors limited capacity of information system users to analyze and use data in planning. In addition to issues discussed previously—such as the lack of training on the collection, analysis, and presentation of data or the lack of appropriate equipment to support data analysis—guidelines or systems were lacking on how to use routine data for decision making. In Tanzania, data use at the facility level was rare owing to a lack of perceived data ownership. Health providers expressed the belief that data could not be used at the point of creation and that they should only concern themselves with data collection. This finding was also seen in South Sudan and Nigeria, where data appeared to be used only to fulfill reporting requirements, not for analysis or
decision making. To encourage data ownership and use at the facility level, one study recommended that supervisors at the district level provide regular feedback to facilities on their data, help facilities analyze the data for their needs, and give providers the opportunity to explain the data at meetings.25

Poor data quality was a barrier to data use for planning and budgeting in multiple studies.18,25,26 Tanzanian researchers found that data quality assurance, particularly accuracy, was a major challenge in the health facilities visited.25 In an in-depth interview, a service provider in Tanzania explained the consequences of poor data quality on decision making as follows25:

*In fact, the work plan is not realistic, there is a big difference between the work plan and budget. As you can see, this center is in the central part of the town. We serve more people than anticipated. For example, the budget has been prepared for 3,880 clients, but we serve 10,000 clients. We normally claim for the same, but they ignore us because we don’t have data. That’s why I say that there is a big difference between work plan and budget; the main reason for this is lack of correct data.* (Service provider at public health facility)

In South Sudan, researchers found that only one health facility included in the study made action-oriented decisions to mobilize or shift resources based on a comparison of services, and only one health facility made evidence-based decisions to advocate for more resources by showing gaps in its ability to meet monthly or annual targets.15 Several studies recommended in-service training to improve providers’ appreciation of how data could inform decisions and build capacity to analyze and use data.14,15,25–27

### DISCUSSION

The findings from the small grant–funded research reports provide an opportunity to identify specific examples of how information system challenges and shortfalls affect data quality and use. Similar to what has been reported in other countries, several small grant–funded studies revealed ongoing challenges with the technology and infrastructure necessary for electronic data collection and reporting.30,31 Although health service providers in multiple study countries expressed overall positive attitudes toward electronic data management and DHIS2, the lack of such basic inputs as providers trained in electronic data capture, a consistent power supply, reliable Internet connectivity, and a sufficient number of operative computers and mobile devices compromises the functionality of RHISs and the success of electronic HISs, including DHIS2. Such difficulties are not specific to family planning; they affect routine health information across all health areas.32 Government investments in these areas will improve the quality and utility of data infrastructure to strengthen the capacity of data management systems at health facilities.

Because many countries’ HISs have been strengthened to capture data on infectious diseases such as HIV, malaria, and tuberculosis, family planning appears to be an afterthought, with less attention and strategic planning for routine family planning data collection and use.22 The successful integration of family planning data in RHISs must accommodate data from disparate sources, ideally through standardized indicators and appropriate use of existing data collection tools along consistent operational guidelines. These tools include patient registries and reporting forms at the clinic, subnational, and national levels, among others. When data are not fully captured and aggregated from all family planning service delivery points and levels in the data management system—as the findings discussed here revealed—they provide an incomplete picture of the status of family planning service delivery and use in a given country. This situation in turn makes evidence-informed decision making difficult. The findings from the research projects pointed to several challenges with data collection tools (e.g., missing forms, incorrect versions, broken mobile devices, lack of guidelines for data collection), human resources (e.g., staff shortages, lack of data management training for personnel, absence of supportive supervision), and governance (e.g., lack of policies and guidelines for submission of data into the national HIS and lack of accountability mechanisms), which also affect data integration and compromise data quality.

Data quality, as defined by data accuracy, relevance, reliability, and timeliness, was found to be problematic in most of the small grant–funded research. Yet each of these characteristics is necessary to ensure integrity of data for policy and programmatic decision making.1,8 A common theme in the research studies was a lack of data training or solid understanding of the HIS and its potential for family planning data analysis and use. This translates into a lack of appreciation for complete, high-quality health data for decision making.

Data as a driver for decision making are integral to HIS performance and the improvement of health systems and outcomes; data use informs...
funding, policy, and national health goals. But if technical, management, organizational, financial, and political barriers to analyzing and using family planning data for planning purposes are present, as was demonstrated across several research studies, initiatives to improve the quality of family planning data will fail to achieve their potential. Fundamental changes in data culture will require strategies to motivate, mentor, and supervise staff at all levels, and staff must be included in programmatic reviews and decisions.

Strengths and Limitations
This synthesis presents the key findings from a body of research produced by local researchers in LMICs supported through MEASURE Evaluation’s 5-year small grants program. The synthesis provides access to research not available through peer-reviewed journals, highlighting context-specific findings from local researchers with specific insight on routine family planning data issues. The research findings have a unique focus on family planning in RHISs, and together provide information about RHISs that is relevant across systems and health areas and specific to the field of family planning. With a focus on routine data (i.e., service statistics), this synthesis identifies several areas for action and intervention to improve the functioning of RHISs and production of reliable, usable family planning information. The synthesis does not, however, attempt to present a comprehensive review of literature on RHISs or family planning information. Furthermore, the identification of key findings and the development of themes are based on the coauthors’ understanding and interpretation of the research. The authors acknowledge that the interconnected nature of routine data capture and production, reporting, analysis, and use make hard boundaries between themes difficult to define. The small grant-funded papers present additional detailed, context-specific research results.

CONCLUSION
The breadth of the small grant-funded research papers revealed several opportunities and barriers related to the integration of family planning data in RHISs in LMICs and the countries’ ability to analyze and use the data to make programmatic and policy decisions. Lack of functioning electronic tools and resources in many contexts prevents providers from fully transitioning to an electronic HIS. A common theme among the study findings was poor data quality resulting from incomplete or missing data from private and nongovernmental organization facilities, insufficient or outdated data collection tools and forms, missing data collection guidelines, poorly defined indicators, and shortages of well-trained data-oriented service providers. Poor-quality data and a lack of data ownership, analysis skills, analysis tools, and a mandate and instruction from higher levels have prevented service providers from learning from their family planning data and making action-oriented decisions. The issues that contribute to poor data quality and its consequences are circular, self-reinforcing, and systemic. Addressing them requires long-term, multipronged interventions to improve family planning data management for well-informed decision making.

Acknowledgments: We wish to express gratitude to the finance and administration staff at Palladium (one of MEASURE Evaluation’s partners) who executed the subgrantee subagreements; the MEASURE Evaluation finance and administration staff at the University of North Carolina at Chapel Hill who provided supporting documentation for the subgranting process; MEASURE Evaluation’s knowledge management team, particularly William Frazier, who expertly edited the subgrantee research manuscripts and served as a writing mentor for the subgrantees; and most importantly, to our 19 subgrantees, who worked to complete successful research projects and add to our knowledge on family planning and routine health information systems.

Funding: This study was made possible by the support of the American People through the United States Agency for International Development (USAID) under the terms of MEASURE Evaluation cooperative agreement AID-OAA-L-14-00004. General support was provided by the Carolina Population Center. The views expressed in this article do not necessarily reflect the views of USAID or the United States Government.

Competing interests: None declared.

REFERENCES


Effectiveness of mHealth Interventions for Improving Contraceptive Use in Low- and Middle-Income Countries: A Systematic Review

Banyar Aung, Jason W. Mitchell, Kathryn L. Braun

Key Findings
- Of the 8 mHealth family planning interventions that met inclusion criteria, 3 studies improved family planning outcomes and 4 studies experienced implementation issues.
- Further research is needed to encourage robust program fidelity of mHealth family planning interventions, along with a more thorough understanding of what mHealth and behavior change components are needed to improve family planning outcomes in low- and middle-income countries.

Key Implications
- A “push” approach, interactive communication, information tailored to participants, motivational messaging, and male partner involvement appear to be tied to better family planning outcomes.
- Program managers and researchers should consider improvements in protocols and fidelity that are needed to more accurately assess how well mHealth family planning interventions impact outcomes in low- and middle-income countries.

ABSTRACT

Background: mHealth interventions are being tested to improve contraceptive uptake in low- and middle-income countries (LMICs); however, the effectiveness of these interventions has not been systematically reviewed.

Objectives: The primary objective of this systematic review was to assess the effectiveness of mHealth interventions to improve contraceptive uptake and adherence in LMICs. A second objective was to identify mHealth features and behavior change communication components used in these mHealth interventions.

Methods: A systematic search was conducted of online databases for peer-reviewed articles that reported on intervention studies with men and women from LMICs and measured mHealth intervention impact on contraceptive uptake and/or adherence. Key search terms included “mHealth” or “mobile health,” “contraception” or “family planning,” and “low- and middle-income countries.” PRISMA guidelines were followed for reporting review methods and findings. The Cochrane risk-of-bias 2 tool for randomized trials was used to assess the risk of bias of the included studies. The GRADE approach was used to determine the quality of evidence.

Results: Eight randomized controlled trial studies met the inclusion criteria. Four studies experienced implementation challenges (e.g., intervention components were not utilized fully by participants, intervention participants did not receive the full intervention content, contamination, low response rate, and/or missing data). Only 3 interventions were found to be effective, and these included a “push” approach, interactive communication, information tailored to participants, motivational messaging, and male partner involvement.

Conclusion: To date, the delivery of mHealth interventions for improving family planning in LMICs has met with implementation challenges that have reduced the researcher’s ability to test intervention effectiveness. Although 3 of 8 studies found improved contraceptive use in the intervention group, the review cannot draw concrete conclusions on the overall effectiveness of mHealth interventions to increase contraceptive use in LMICs. Further research with robust program fidelity is recommended.

INTRODUCTION

By the end of the Millennium Development Goals in 2015, the maternal mortality ratio had declined by 45% from 1990.1 Despite this progress, every day, 810 mothers—94% from low- and middle-income countries
Most mHealth research on family planning and uptake of modern contraceptives has occurred in higher-income countries, with few trials and studies in LMICs.

(LMICs)—continue to die from preventable causes associated with pregnancy and childbirth. To help reduce maternal deaths in LMICs, the Safe Motherhood Initiative outlined family planning as 1 of 6 “pillars” of safe motherhood. One viable solution to reduce maternal mortality in LMICs is to increase the uptake of contraceptives, which in turn will reduce the number of unwanted pregnancies. Fulfilling the current unmet need for contraceptives has occurred in higher-income countries, with few trials and studies having occurred in LMICs.

In sum, the bulk of research involving mHealth on reproductive health and family planning has been conducted in the United States, Cambodia, and Israel. Only one of the studies occurred in an LMIC. The review concluded that interactive voice messages and communication with a counselor improved postabortion contraception, and the combination of unidirectional (i.e., one-way messages) and interactive daily educational text messages (i.e., back-and-forth) improved adherence in using oral contraceptives.

In another review, L’Engle et al. examined 35 studies that used mobile phones to improve adolescent sexual and reproductive health, inclusive of contraceptives. Only 3 of the 35 studies were from LMICs. The authors found evidence that including text messages in interventions may improve adolescent sexual health, but the information provided in the studies was insufficient for understanding, replicating, or scaling up mHealth interventions.

Rousseau et al. conducted a systematic review with 22 studies to explore the general impact of smartphone applications on contraceptive decision making and knowledge. Fifteen of the 22 studies were based in the United States, 3 were conducted in an LMIC, and the locations of the 4 remaining studies were not specified. The reviewers found that apps may be useful as aids to improve contraceptive use and prescription of contraception, but they were not reliable sources of information. The authors noted that the quality of the studies was heterogeneous, adding to the difficulty in drawing conclusions about the impact of mHealth apps on contraceptive knowledge and usage.

Although previous systematic reviews assessed the effectiveness of mHealth interventions for family planning, only the review by Smith and colleagues focused exclusively on contraceptive uptake, while other 2 systematic reviews involved other outcomes (e.g., contraceptive knowledge). Furthermore, only 7 studies included in these 3 reviews were based in an LMIC, and only 1 measured contraceptive use. In sum, the bulk of research involving mHealth on family planning and uptake of modern contraceptives has occurred in higher-income countries, with few trials and studies having occurred in LMICs. Given the disparities of maternal mortality and unmet family planning needs in LMICs, a more

Three published reviews explored the effectiveness of mHealth interventions for different contraceptive outcomes. Smith et al. assessed the effect of interventions delivered via mobile phone for improving contraceptive use in 5 randomized controlled trials (RCTs) conducted in the United States, Cambodia, and Israel. Only one of the studies occurred in an LMIC. The review concluded that interactive voice messages and communication with a counselor improved postabortion contraception, and the combination of unidirectional (i.e., one-way messages) and interactive daily educational text messages (i.e., back-and-forth) improved adherence in using oral contraceptives.

In another review, L’Engle et al. examined 35 studies that used mobile phones to improve adolescent sexual and reproductive health, inclusive of contraceptives. Only 3 of the 35 studies were from LMICs. The authors found evidence that including text messages in interventions may improve adolescent sexual health, but the information provided in the studies was insufficient for understanding, replicating, or scaling up mHealth interventions.
thorough examination of the role of mHealth in improving the uptake of modern contraceptives in LMICs is needed.

The primary objective of the present systematic review was to assess the effectiveness of mHealth interventions in improving contraceptive uptake in LMICs. The secondary objective of the systematic review was to identify which mHealth features and behavior change communication (BCC) components were used in the mHealth interventions that occurred in LMICs.

**METHODS**

Review findings are reported based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses-Protocols (PRISMA) guidelines. The review protocol was preregistered in the PROSPERO database (CRD42020153409).

**Inclusion Criteria**

**Type of Studies**

Experimental studies that evaluated the intervention effectiveness through RCTs and nonrandomized interventional studies were considered for the review.

**Type of Participants**

Women and men from LMICs, as classified by the World Bank, were included. The WHO definition of women of reproductive age (15–49 years old) was not used because more than 1 in 3 (about 250 million) girls were married or in union before age 15, with the highest rates found in LMICs in South Asia and sub-Saharan Africa. Postpartum and postabortion women were also included.

**Type of Interventions**

We included studies in which the intervention was delivered using any form of mHealth such as mobile apps, messaging platforms or short messaging system (SMS), telephone calls, or geolocational features (e.g., GPS or Global Positioning System). We included the interventions that sought to improve contraceptive uptake and/or adherence compared with standard care or another intervention. mHealth interventions were identified based on the definition of the WHO Global Observatory for eHealth.

**Type of Outcome Measures**

For the purposes of this review, we included the outcome measurement of uptake or adherence to any modern contraceptives including permanent methods (female sterilization and vasectomy); long-acting reversible contraceptives (implants and intrauterine devices); and shorter-acting contraceptives (injectables, pills, male and female condoms, diaphragms, spermicides, and cervical caps). We acknowledge that other nonbiomedical methods such as fertility awareness methods and withdrawal methods exist, but these were not included in our definition. We accepted whichever method by which the outcome was assessed in the included mHealth intervention trials/studies, including by self-report through surveys. Interventions were included even if the uptake and/or adherence to contraception was not the primary outcome measured or was measured in conjunction with other contraceptive outcomes such as knowledge of contraception.

**Search Strategy**

The search was conducted by BA in July 2019. A filter was set to include articles from 2005, since mobile subscriptions reached 23% of populations in LMICs in 2005 (compared with only 4% in 2000). PubMed, Web of Science, EBSCOhost, CINAHL, and The Cochrane Library were searched. Key search terms used were “intervention”; “program”; “mHealth”; “mobile health”; “teledmedicine”; “cell phone”; “SMS”; “apps”; “contraception”; “contraceptive”; “family planning”; “birth spacing”; “developing country”; and “low and middle income country”. LMICs were further searched by detailing regions such as Africa, Asia, Pacific Islands, South America, Central America, Latin America, Eastern Europe, and Central Asia. Reference lists of identified articles were searched. We retrieved study protocols of included studies and assessed method details. We contacted authors of included studies if the study protocol was not published and when additional information was needed. Only articles in English were included due to the reviewers’ language limitation.

**Data Collection and Analysis**

The search was completed by BA and KLB independently. Duplicates were removed and titles and abstracts were assessed applying the inclusion criteria. Screened articles were read in full, again, by applying the inclusion criteria. Discrepancies were resolved by discussion.

BA and KLB extracted information from the studies, including the country in which the study was conducted, intervention details (e.g., mHealth features, mode of delivery, BCC components, frequency, duration), participant characteristics (age,
We did not conduct a meta-analysis due to the diversity of intervention components and outcome measures that were used in the included studies. However, clinical heterogeneity (i.e., variability in participants, interventions, outcomes studied) and methodological heterogeneity (i.e., variability in study design and risk of bias) of the included studies were characterized.

Assessment of Publication Bias
We were unable to execute a funnel plot to identify the publication bias due to the diversity of intervention components and outcome measures that were used in the included studies.

Data Synthesis
We conducted the analysis according to the guidelines specified in the Cochrane Handbook for Systematic Reviews of Interventions.37 Quality of evidence for included studies was assessed using the Grading of Recommendations Assessment, Development and Evaluation Working Group (GRADE) approach.38 RCTs were considered high quality and were downgraded by 1 level for “serious” (or 2 levels for “very serious”) risk of bias; unexplained heterogeneity; indirectness of evidence; imprecision of effect estimates; or publication bias.

RESULTS
Figure 1 shows the PRISMA flow diagram for the systematic review. Among the 123 publications identified in the database search, 43 duplicates were removed and 80 studies were assessed; all 80 studies were published in English. After titles and abstracts were screened, 73 studies were excluded for not meeting inclusion criteria and 7 articles were further assessed. One additional article was identified through reference tracing. Eight studies met the inclusion criteria for this systematic review.

Study Characteristics
Of the 8 studies included in this review, 3 were conducted in Kenya,18,23,25 1 in Cambodia,22 1 in Ecuador,19 1 in Tajikistan,20 1 in Palestine,21 and 1 in Bangladesh.24 (Table 1). All 8 studies were parallel-group RCTs with 1:1 allocation, including a feasibility study with a small sample size.24 Study settings varied. Some were conducted in urban or peri-urban and rural areas,22 while others were conducted in a hospital or clinic setting.18,23–25; settings for 2 studies were not specified.20,21 The studies also varied by types of participants: postpartum mothers,19,23,25 postabortion women,22,24 young people,20,21 and general public.18 Outcomes for 6 of 816
the 8 studies were about contraceptive use and knowledge,18,20,22,24,25; the other 2 studies measured the same outcomes and other maternal and child health indicators (e.g., exclusive breastfeeding and immunization coverage).19,23 Three of the 8 studies provided a description about the use of behavior change theory.20,21,25

mHealth Features

mHealth features used in the 8 included studies varied (Table 2). Two studies used telephone calls. Smith et al.22 delivered interactive voice messages and provided counselor support via telephone calls upon participants’ request through the messages. Counselor phone support involved tailored information a range of contraceptive methods and motivation about using contraception, as well as helping participants in their search for family planning clinics. In contrast, the telephone call was made by a nurse to deliver health education about family planning in the study by Maslowsky et al.19

Six studies used text messages as their primary mHealth feature to deliver health education and motivational messages about family planning.18,20,21,23–25 McCarthy et al.20 included an app in their intervention in Tajikistan to mainly deliver one-way text messages about contraception, common beliefs on family planning, and encouragement to use family planning. A similar intervention content was delivered via one-way text messages (without an app) to participants in the study conducted in Palestine by McCarthy et al.21 Using a text messaging platform named m4RH, Johnson and colleagues delivered information about family planning, a searchable database of clinics providing family planning services, and an optional role model stories feature.18

Mobile SMS delivery platform Mobile WACh and its variant Mobile WACh XY (with male partner involvement) were used by Unger et al.23 and Harrington et al.,25 respectively, to provide interactive intervention contents tailored to participant needs. The intervention tested by Biswas et al.,24 in Bangladesh, was a feasibility study conducted with a small sample size and it found no effect. Method-specific text message reminders were sent to participants about their select methods. It only involved unidirectional SMS reminders without other BCC components. However, the study found an mHealth contraceptive intervention was feasible, citing positive user engagement and participant acceptability.

Interactive communication was used in 4 studies,19,22,23,25 Seven studies used a push approach whereas only 1 study used a pull approach18 to deliver intervention content to participants. Of the 3 studies that reported improving contraceptive uptake,22,23,25 all used a
push approach to deliver information and an interactive type communication.

**BCC Components**

Interventions utilized different intervention components to facilitate behavior change (Table 2), ranging from motivation to use family planning,20–23 tailoring of information,19–23 partner involvement,22,25 service provider search features,18,22 and role model stories.18 Two of the 3 interventions that reported improved contraceptive uptake included the involvement of a voluntary male partner.22,25 The study by Smith et al.22 involved counselor phone support that was tailored to the participant’s need and provided motivation to use postabortion contraception and information on nearest service providers. The intervention by Unger et al.23 provided weekly unidirectional (partial intervention) and interactive (full intervention) family planning related educational and motivational SMS tailored to the recipient, and found that both full and partial interventions improved early postpartum contraceptive use over the control condition. The

### TABLE 1. Summary of Studies Included in Systematic Review of mHealth Interventions Assessing Contraceptive Uptake in Low- and Middle-Income Countries, N=8

<table>
<thead>
<tr>
<th>Authors</th>
<th>Country</th>
<th>mHealth Delivery Mode</th>
<th>Target Population</th>
<th>Study Design</th>
<th>Sample Sizea (Intervention/Control)</th>
<th>Frequency and Duration</th>
<th>Posttest and Follow-Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Johnson et al.18</td>
<td>Kenya</td>
<td>HE via text messaging, “role model” stories, clinic database</td>
<td>General public</td>
<td>RCT (probably unblinded)</td>
<td>13,629 (6,817/6,812)</td>
<td>Over 3 months</td>
<td>24 hours, 6 days, 3 months postenrollment</td>
</tr>
<tr>
<td>Maslowsky et al.19</td>
<td>Ecuador</td>
<td>Telephone-delivered HE and telephone access to a nurse</td>
<td>Postpartum women</td>
<td>Unblinded RCT</td>
<td>178 (102/76)</td>
<td>Within 48 hours of hospital discharge. Access to a nurse on-call during the first 30 days of the newborn’s life</td>
<td>3 months after delivery</td>
</tr>
<tr>
<td>McCarthy et al.20</td>
<td>Tajikistan</td>
<td>HE via app instant messaging</td>
<td>Young people (16–24), both genders</td>
<td>Single-blinded RCT</td>
<td>543 (275/298)</td>
<td>0–3 messages per day over 4 months</td>
<td>4 months after baseline</td>
</tr>
<tr>
<td>McCarthy et al.21</td>
<td>Palestine</td>
<td>HE via text messaging</td>
<td>Young women (18–24)</td>
<td>Single-blinded RCT</td>
<td>578 (289/289)</td>
<td>0–3 messages per day over 4 months</td>
<td>4 months after baseline</td>
</tr>
<tr>
<td>Smith et al.22</td>
<td>Cambodia</td>
<td>Voice messages and phone calls</td>
<td>Women, postabortion</td>
<td>Single-blinded RCT</td>
<td>300 (249/251)</td>
<td>6 automated voice messages ± telephone counseling within 3 month period</td>
<td>4 and 12 months postabortion</td>
</tr>
<tr>
<td>Unger et al.23</td>
<td>Kenya</td>
<td>HE via text messaging</td>
<td>Postpartum women</td>
<td>3-arm, unblinded RCT</td>
<td>300 (100/100/100)</td>
<td>Weekly until 12 weeks postpartum</td>
<td>From antenatal care attendance and followed through 10, 16, 24 weeks postpartum</td>
</tr>
<tr>
<td>Biswas et al.24</td>
<td>Bangladesh</td>
<td>HE via text messaging</td>
<td>Women, postabortion</td>
<td>RCT (probably unblinded)</td>
<td>120 (60 /60)</td>
<td>Method-specific reminders/intervals (daily/weekly)</td>
<td>4 months postabortion</td>
</tr>
<tr>
<td>Harrington et al.25</td>
<td>Kenya</td>
<td>HE via text messaging</td>
<td>Postpartum women</td>
<td>Unblinded RCT</td>
<td>254 (125/129)</td>
<td>Weekly from enrollment to 6 months postpartum</td>
<td>6 months postpartum</td>
</tr>
</tbody>
</table>

Abbreviations: HE, health education (contraceptive information); RCT, randomized control trial; app, mobile application.

a Data from participants who were analyzed.
b Authors did not mention about blinding. This information was deduced from reading the studies.
study by Harrington et al. used a variant of the intervention used by Unger et al. (Mobile WACH), but with a voluntary male partner involvement (Mobile WACH XY). However, male involvement did not have a significant effect on contraceptive use outcomes compared with having women as the only participants. In terms of frequency of intervention delivery, findings from these studies suggest that improved contraceptive use was associated with weekly or biweekly messaging rather than daily or a one-time delivery.

**Study Quality and Risk of Bias: Cochrane’s RoB 2 Tool**

**Randomization Process**

Cochrane’s RoB 2 tool (Figure 2) classified 5 studies as low risk and 3 studies with some concerns for risk of bias in the randomization process domain. Johnson and colleagues used the alternation method of allocating participants to intervention and control groups, instead of true randomization. The trials by Harrington et al. and Biswas et al. had important baseline differences between their control and intervention groups.
Deviation From Intended Interventions

Five studies were conducted as intended and were thereby classified as having a low risk of bias (Figure 2). However, 3 studies deviated from their original protocol, resulting in a high risk of bias for this domain. The intervention by Maslowsky et al. had 2 parts, yet only 1 was delivered. It also had substantial contamination, with controls also receiving the intervention, as occurred in the intervention of McCarthy et al. In another study, participants did not receive the complete intervention.

Missing Outcome Data

Seven studies had low risk of bias for missing outcome data (Figure 2). The study by Johnson et al. had some concerns for risk of bias due to having low retention rates: 20.9% of intervention and 21.3% of control participants were lost to follow-up for surveys that measured contraceptive uptake. To overcome this problem, researchers used multiple imputation methods for both groups but some concerns remain for risk of bias in this domain.

Measurement of Outcome

All 8 studies had high risk of bias for measurement of the outcome (Figure 2). All the studies relied on self-reported outcomes obtained from final assessments; thus, the assessors were the participants and the outcome measurement may have been subjected to social desirability bias. The collection of outcome data was not blinded.

Selection of the Reported Result

As shown in Figure 2, half of the studies were at low risk for selective outcome reporting since all outcomes were reported in their results. The other half of studies had some concerns for selection of the reported result because the protocols containing details about their prespecified analytic plan were not published.

Overall Risk of Bias

Cochrane’s RoB 2 tool classifies the overall risk of bias to be considered high risk if any of individual domains (e.g., randomization process, missing outcome data) assessed were deemed high risk. As a result, all 8 studies were labeled as having an overall high risk of bias. Figure 3 provides a

FIGURE 2. Risk of Bias in Studies of mHealth Interventions to Increase Contraceptive Uptake in Low- and Middle-Income Countries, N=8
DISCUSSION

Summary of Findings

To our knowledge, this systematic review is the first to assess the effectiveness of mHealth interventions toward increasing contraceptive use in LMICs. Other systematic reviews have examined mHealth in family planning interventions, but few included studies from LMICs. Additionally, 7 of the 8 studies in the present review were not assessed in previous systematic reviews. Findings from the current systematic review reveal new information about the role that mHealth and BCC components have in improving contraceptive use in LMICs.

Of the 8 included studies, 3 reported improvements in family planning outcomes among people who received the intervention compared with controls. With respect to mHealth, 2 of the 3 studies used text messages, while the other study used voice messages and telephone counseling, which included information about the nearest family planning service provider. Two common traits that the 3 studies shared were the use of interactive communication and a push approach to deliver tailored intervention content to participants. Other commonalities were the use of motivational messages and the involvement of a male partner in the intervention.

Given that only 3 of the 8 studies found improvements in family planning outcomes, the full extent that mHealth contributed to improvements in the use of modern contraceptives among participants cannot be determined. It is possible that certain types of mHealth features may be more advantageous to effect change in the use of modern contraceptives. For example, interactive communication and the use of a push approach to deliver intervention content entails engagement with participants. The frequency that the intervention information is delivered in studies that used the push approach may also have an impact on participants’ use of modern contraceptives. Some studies in this review delivered intervention information once, daily, weekly, or biweekly. Positive changes in outcomes were found in studies that delivered the intervention information on a weekly or biweekly basis, suggesting too frequent delivery may not resonate with participants with respect to their family planning needs.

Analysis of the BCC components used among the 8 included studies suggests tailoring information to the participant and potentially the use of motivational messages and/or the involvement of a male partner may play a role in improving contraceptive use. Among the 3 studies that showed significant improvements in outcomes (intervention vs. control), all tailored the information delivered, whereas 2 of the studies used motivational messages and 2 involved the male partners of participants. However, Harrington et al. conducted a subgroup analysis...
Interventions improving contraceptive uptake combined unidirectional and interactive communication styles and used multiple BCC components.

Behavioral change theories are important in improving targeted behaviors, and further investigation will identify which mHealth and BCC components lead to better outcomes.

Comparison With Existing Literature

Our review found that interventions that showed significant improvement in contraceptive uptake used a combination of unidirectional and interactive communication styles and involved multiple BCC components. Notably, simple unidirectional text message reminders had no effect on improving contraceptive uptake. Such findings are consistent with the evidence from the systematic review that assessed the effect of mHealth interventions to improve contraceptive uptake, with 80% of studies involved having been conducted in developed countries.

The International Conference on Population and Development set the involvement of men in family planning as a priority area. They are similar derivative theories associated with fertility decline in the family (due to increased use of contraception) and resulted in long-term benefits for women. For example, a case study spanning 5 generations of a family in an LMIC setting found that male involvement in family planning was associated with fertility decline in the family and a decreasing trend in family size. In another study, Tao et al. found that involvement of the male partner in family planning decision making improved family planning knowledge and contraceptive continuation. Moreover, a systematic review that examined different BCC techniques used to improve contraceptive use in LMICs found that the most effective interventions were those that involved male partners. Prior research suggests the involvement of male partners differs (e.g., sexual/romantic relationship, family, friend), as well as the amount and frequency of their involvement toward achieving these outcomes.

Only 3 studies included in this review reported using a behavioral change theory. Two of them were conducted by the same researchers, who used the Integrated Behavioral Model and the other study used the Theory of Planned Behavior. They are similar derivative theories of general behavioral prediction, with the most important determinant being motivation or intention as the interventions targeted. A systematic review by Cho et al. examined the use of theories in mHealth behavior change interventions conducted in the LMICs and also found that about one-third (5 of 14) of their included studies were based on a behavioral change theory. Well-tested behavioral change theories are useful to help guide the design and implementation of family planning interventions and programs. As the effectiveness of mHealth in family planning interventions in LMICs remains inconclusive, future research that uses behavioral change theory for contraception uptake is warranted and needed to help identify which intervention components (mHealth and behavior change) work best for family planning and why.

Systematic reviews on behavior change interventions of other health topics that used mHealth recommended the inclusion of certain components to increase the effectiveness of the intervention. For example, a systematic review on technologically driven weight-loss interventions by Khaylis et al. identified the following components as essential for improving outcomes: use of behavior change theory, self-monitoring, counselor feedback and communication, social support (motivation), and tailoring information. A meta-analysis by Webb et al. recommended that technology-based interventions make extensive use of theory, incorporate more BCC techniques, and use SMS or text messages to effectively promote behavior change. These reviews, along with the present one, suggest that the use of behavioral change theories is important to improve targeted behaviors, while also recognizing that further investigation is warranted to decipher which mHealth and BCC components and in what combinations lead to better family planning outcomes.

Considerations of Intervention Fidelity, Missing Data, and Limited Use

As noted in the risk-of-bias assessments, some studies included in this review reported issues with intervention fidelity or missing data. Findings from this review found important shortcomings in the included interventions that may have affected the study’s findings. Four out of 5 studies that did not find any significant changes in outcomes between trial arms had poor implementation or retention issues.

Regarding fidelity, the study conducted in Tajikistan by McCarthy et al. found contamination between trial arms (i.e., some controls received...
a portion of the intervention content) because of a misunderstanding between research partners. As such, the trial was assessed as the full intervention versus the partial intervention, instead of what was originally planned (i.e., comparing between intervention and control). Another study by McCarthy and colleagues\(^{21}\) conducted in Palestine had technical problems with the messaging platform used, which resulted in 60% of the intervention participants not receiving the full intervention. The outcomes measured were only based on the effect of partial receipt of the intervention versus the control. Further, contamination may have also occurred: 17% (39/235) of the control participants reported reading messages for someone else in the study and 17% (40/229) of intervention participants said that someone else in the study read their messages.

With respect to missing data, the intervention tested by Johnson et al.\(^{18}\) offered new users the m4RH app with text-message-based family planning information as well as a searchable database of service providers, with an option to receive role model stories of current users. The study had low response rates to its 3 assessments (range: 51.8% to 13.5%), and the proportion of participants who responded to more than 1 assessment was low. This large number of missing longitudinal data affected the statistical power for the study analyses, which may have influenced their findings.\(^{18}\)

Not all participants will use all parts of an intervention. For example, Maslowsky et al.\(^{19}\) designed a 2-part intervention, with part 1 consisting of a one-time telephone-delivered health education session and part 2 consisted of having access to an on-call nurse for personalized advice (via telephone). Of the 178 study participants, only 3 participants used part 2; participants had to take the initiative to use part 2. Access to the on-call nurse for personalized advice included motivational support and tailored information, including where to receive contraceptive services. Numerous reasons may exist for why part 2 of the intervention was not used by the study participants and how its use and nonuse may have impacted the study’s

---

### TABLE 3. Quality of Evidence of the Contraceptive Uptake Outcome Using the GRADE Approach in Studies Included in the Review, N=8

<table>
<thead>
<tr>
<th>Study</th>
<th>Limitations of Detailed Design and Execution (Risk of Bias)</th>
<th>Unexplained Heterogeneity or Inconsistency of Results</th>
<th>Indirectness of Evidence</th>
<th>Imprecision of Results</th>
<th>Publication Bias</th>
<th>Quality of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smith et al.(^{22})</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Maslowsky et al.(^{19})</td>
<td>-2</td>
<td></td>
<td>-1</td>
<td></td>
<td></td>
<td>++++ Very low</td>
</tr>
<tr>
<td>McCarthy et al.(^{20})</td>
<td>-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++++ Low</td>
</tr>
<tr>
<td>Johnson et al.(^{18})</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++ Moderate</td>
</tr>
<tr>
<td>Unger et al.(^{23})</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++ Moderate</td>
</tr>
<tr>
<td>McCarthy et al.(^{21})</td>
<td>-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++ Low</td>
</tr>
<tr>
<td>Biswas et al.(^{24})</td>
<td>-1</td>
<td></td>
<td>-1</td>
<td></td>
<td></td>
<td>++ Low</td>
</tr>
<tr>
<td>Harrington et al.(^{25})</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++ Moderate</td>
</tr>
</tbody>
</table>

Randomized controlled trials were considered to be high quality, but were downgraded by 1 level (serious) or 2 levels (very serious) for each of the following: limitations of detailed design and execution (risk of bias) (e.g., limitations in randomization, deviations from intended interventions), unexplained heterogeneity or inconsistency of results, indirectness of evidence, imprecision of results, and presence of publication bias.
findings. Future mHealth interventions for family planning ought to integrate monitoring of intervention delivery and other process evaluation techniques, as well as brief qualitative exit interviews or quantitative measures (e.g., Health ITUES), \textsuperscript{31} to better understand the reasons why participants use and do not use certain parts of an intervention and how their usage affects the study’s findings.

Successful intervention outcomes necessitate well-implemented programs, and implementation fidelity is crucial for the intervention effectiveness. \textsuperscript{52} Half of the studies included in this review reported poor implementation or retention issues, which limits the ability to fully evaluate the intervention and assess its impact on contraceptive uptake outcomes. Future mHealth family planning trials ought to implement steps to help ensure the fidelity to the protocol and design of the intervention.

Quality of the Evidence

Quality of the evidence was assessed using the GRADE approach \textsuperscript{38} (Table 3). Five trials were downgraded by 1 level under the domain of limitations in design and execution because they both had a high risk of bias in the measurement of the outcome. \textsuperscript{22,23} Under the same domain, 3 trials \textsuperscript{19–21} were downgraded by 2 levels due to high risk of bias from deviations from intended intervention, in addition to high risk of bias in measurement of outcome. Two trials were downgraded by 1 level under the imprecision of results domain due to small sample sizes. \textsuperscript{19,24} Overall, the quality of evidence was graded as moderate in 4 trials, low in 3 trials, and very low in 1 trial.

Self-reported outcomes are the standard in contraceptive research, but they are subject to social desirability bias. \textsuperscript{53} Additionally, intervention and control participants recruited from the same hospital or clinics might have shared intervention contents with each other, resulting in contamination.

Limitations

It is important to acknowledge that this review only included RCTs and nonrandomized studies to evaluate the effectiveness of mHealth interventions. Other types of evidence may exist and ought to be considered when evaluating the effectiveness of mHealth interventions for family planning. For example, policy makers and other key stakeholders may find equal value from assessing how well an mHealth-mediated family planning program has achieved its goals and outcomes through other types of study designs that blend research with evaluation (e.g., one-group). Another consideration pertains to whether evidence on mHealth interventions conducted in LMICs is disseminated in peer-reviewed outlets (e.g., journals), as noted by Gurman et al. \textsuperscript{54} in their systematic review.

CONCLUSION AND RECOMMENDATIONS

The use of mobile phones and smartphones in LMICs has proliferated, suggesting mHealth might be a viable tool for delivering interventions aimed at improving family planning outcomes. However, there is insufficient evidence to conclude whether mHealth interventions improve contraceptive uptake in LMICs based on the findings from this review and other systematic reviews. \textsuperscript{26,28} Although 3 of 8 studies in this review showed significant improvement in contraceptive outcomes, their effectiveness cannot be linked to specific mHealth features or BCC components.

Moreover, the quality of evidence suggests that improvements in the implementation fidelity and use of behavior change theories are needed for future mHealth family planning interventions in LMICs. Further investigation is warranted to assess and identify which mHealth features, BCC components, and theories, as well as in what specific combinations, will lead to better family planning outcomes and for which specific groups and LMIC locations.

Competing interests: None declared.

REFERENCES


A Practical Guide to Using Time-and-Motion Methods to Monitor Compliance With Hand Hygiene Guidelines: Experience From Tanzanian Labor Wards

Giorgia Gon, Said M. Ali, Robert Aunger, Oona M. Campbell, Micheal de Barra, Marijn de Bruin, Mohammed Juma, Stephen Nash, Amour Tajo, Johanna Westbrook, Susannah Wood, Wendy J. Graham

Key Messages

- Time-and-motion methods are a good way of capturing hand hygiene compliance; for example, they can limit observer bias.
- We describe how we designed the HANDS at Birth tool, the tool format and its elements, its implementation components, the tool’s performance, and the implications for data analysis.
- The advantages of using this tool include simpler training, less observer judgment in assessing hand hygiene compliance, and improved ability to monitor multiple behaviors.

ABSTRACT

Background: Good-quality evidence on hand hygiene compliance among birth attendants in low-resource labor wards is limited. The World Health Organization Hand Hygiene Observation Form is widely used for directly observing behaviors, but it does not support capturing complex patterns of behavior. We developed the HANDS at Birth tool for direct observational studies of complex patterns of hand rubbing/washing, glove use, recontamination, and their determinants among birth attendants. Understanding these behaviors is particularly critical in wards with variable patient volumes or unpredictable patient complications, such as emergency departments, operating wards, or triage and isolation wards during epidemics. Here we provide detailed information on the design and implementation of the HANDS at Birth tool, with a particular focus on low-resource settings. We developed the HANDS at Birth tool from available guidelines, unstructured observation, and iterative refinement based on consultation with collaborators and pilot results. We designed the tool with WOMBAT software, which supports collecting multidimensional time-and-motion data. Our analysis of the tool’s performance centered on interobserver agreement and convergent validity and the implications of the data structure for data analysis. The HANDS at Birth tool encompasses various hand actions and context-relevant information. Hand actions include procedures relevant during labor and delivery; hand hygiene or glove actions; and other types of touch. During field implementation, we used the tool for continuous observation of the birth attendant. Interobserver agreement was good (kappa range: 0.7–0.9), and the tool showed convergent validity. Using the HANDS at Birth tool is a feasible way to obtain useful information about compliance with hand hygiene procedures. The tool could be used after simple training and allows for collection of reliable information about the complex pattern of hygiene behaviors. Future studies should explore using this tool to observe behavior in labor wards in other settings and in other types of wards.

BACKGROUND

Infection prevention is paramount to limiting the spread of epidemics, such as coronavirus disease 2019 (COVID-19), severe acute respiratory syndrome, and Ebola, and hand hygiene (HH) is at the forefront of prevention efforts among health care workers. In addition, health care workers’ HH is essential at the time of birth for preventing
Multiple methods exist to measure HH compliance in health care settings, but observation of behaviors is considered to be the gold standard. Observation can be done by an observer or by video recording. A recent validation study suggests that both approaches capture similar numbers of HH opportunities—moments when health care workers ought to practice hand rubbing/washing; however, video recording poses substantial ethical issues, which often makes it difficult to use, particularly in a process such as childbirth when women are vulnerable and undressed. The World Health Organization (WHO) HH Observation Form is an excellent, widely used tool for direct observation. However, due to its aim and scope, it does not allow capturing more complex patterns of behavior. For example, it does not distinguish whether the failure to comply was because hand rubbing/washing was not attempted or because hands were recontaminated after initial washing. Avoiding hand/glove recontamination is implicit in the WHO tool’s HH definition because touching a surface carries the risk of germ transmission and creates a new HH opportunity. It also does not aim to capture the use or “misuse” of gloves. Finally, it requires the observers to judge when a new HH opportunity arises, thereby reducing the consistency of data collection by multiple observers.

Defining when a new HH opportunity arises is particularly difficult in labor and delivery, during which observers must deal with a transition from observing 1 patient (the mother) to 2 (mother and newborn). Furthermore, the amount, type, and location of body fluids can rapidly change during labor and delivery, and in the context of low-resource settings, a single health care worker may attend many mothers simultaneously. With an often unpredictable duration of the different stages of labor, the time between hand rubbing/washing and delivery of the newborn may be lengthy, during which time the observer needs to pay close attention to assess if any actions occur that lead to a new HH opportunity. Time-and-motion methods can overcome some of these challenges. These methods are now at the forefront of health care observation and are increasingly used, but seldom in low- and middle-income countries. These methods enable observers to record all health care workers’ actions without having to decide which ones represent a new HH opportunity. Instead, opportunities are defined during data analysis.

The HANDS study (Hand-hygiene of Attendants for Newborn Deliveries and Survival) was a mixed-methods, cross-sectional research study conducted in the 10 highest volume maternity wards in Zanzibar between November 2015 and April 2017. The aim of the study was to explore compliance with HH guidelines and identify factors that explain compliance. HH during labor and delivery is a key opportunity to prevent transmitting infections to mothers and newborns; however, good-quality evidence on HH compliance from low-resource labor wards is limited. Therefore, we developed the HANDS tool, based on a time-and-motion design, to observe the complex patterns of birth attendants’ HH and glove use at 3 levels: the opportunity, the individual, and the facility. We designed the tool within WOMBAT software, which is intended to support direct observational studies of health care work. The WOMBAT software package allows collecting multidimensional work tasks, including compliance with specific tasks, and automatically time-stamps data entry. The current investigation was one of the few time-and-motion studies of health care workers conducted with software that automatically records time and carried out in a low-resource setting.

Our aim was to provide very practical details regarding the design and implementation of the direct observational tool to measure HH compliance to inform researchers and practitioners seeking to thoroughly measure the compliance with HH guidelines during labor and delivery, particularly in low-resource settings. In this article, we outline: (1) how we designed the data collection tool, (2) the tool format and its elements, (3) its implementation components, (4) the tool’s performance, and (5) the implications for data analysis.

METHODS FOR TOOL DEVELOPMENT
We developed the HANDS at Birth data collection tool between March and October 2016 using an existing systematic process for tool development. This process included use of available guidelines, unstructured observation, and iterative refinement based on consultation with collaborators and pilot results.

Guidelines’ Review and Semistructured Observation
in Outpatient and Home-Based Care and Long-term Care Facilities, and Pregnancy, Childbirth, Postpartum and Newborn Care: A Guide for Essential Practice. We also conducted 11 semi-structured observation sessions in 4 labor wards in Zanzibar during which either a delivery or a vaginal examination occurred. All birth attendants’ actions were recorded, together with the time when they happened and their location. Using this information, we created a list of procedures (what we also call “key attendant-patient interactions”) relevant to labor and delivery that also included other hand actions that can occur before and after each of these procedures.

Iterative Collaborator Consultation
The project was a partnership of the London School of Hygiene and Tropical Medicine, the University of Aberdeen, and the Public Health Laboratory of Pemba; we sought feedback on the tool from all project members. Additionally, a 3-hour in-depth consultation was conducted with 2 clinically trained members of the team (1 general practitioner and 1 midwife) who provided additional feedback.

Pilot Activities and Training
We conducted 3 pilot activities in a labor ward on Pemba Island, Zanzibar, Tanzania. Two data collectors conducted the first pilot in June 2016 using an early version of the HANDS at Birth tool. One data collector conducted the second pilot in August 2016 using the tool incorporated into WOMBAT v2 software on a tablet. Finally, 1 data collector conducted the third pilot in September 2016 using the tool with WOMBAT. Feedback was collected and incorporated to improve the tool at each stage.

Observers were trained to use the tool over 3 days using role-plays and presentations. Each observer also practiced using the tool in the labor ward for 3 hours under trainer supervision (GG). The trainer also conducted 2 hours of observation with each observer and provided them with relevant feedback. During training, minor refinements were made to the tool.

We used the STROBE checklist for cross-sectional studies to design and describe this tool here and in other relevant manuscripts including the study results.

The project was approved by the Zanzibar Medical Research and Ethics Committee, the London School of Hygiene and Tropical Medicine Research Ethics Committee, and the Research Ethics Committee at the University of Aberdeen. Details of procedures to consent are described below.

TOOL FORMAT AND ELEMENTS
Following Lopetegui et al.’s classification, our time-and-motion study used continuous observation, in which an external observer focuses on 1 subject, in our case, the birth attendant. When a birth attendant performed an action, the observer recorded the action. We chose to use continuous observation because the timing of procedures, particularly delivery itself, was typically unpredictable, and using alternative methods, such as short observation sessions at fixed or random intervals, could have missed many HH opportunities. Hence, observers were asked to remain in the labor room for the entirety of their allocated shift (about 7 hours for morning/afternoon shifts and 10 hours for night shifts) and to start recording observations whenever a patient-attendant interaction began.

The tool, available in Supplement 1, includes a list of hand actions and context-relevant information (Figure). The hand actions listed were exhaustive (meaning that the list did not leave any possible actions out) and mutually exclusive (meaning that no 2 actions could occur simultaneously). We did not design a tool that aimed to capture multitasking or interruptions because we did not want to add to the burden on the observers.

Hand actions were either procedures relevant during labor and delivery (e.g., vaginal examination) (Table 1), HH or glove actions, or other types of touches (e.g., touching a pen or equipment). Observers recorded when an attendant left the room where observation was occurring (when observation was suspended) and when the attendant re-entered.

The tool also captured information on the context, such as availability of key infrastructure/staffing (e.g., water or the presence of the nurse in-charge) and which woman was being attended (first, second, third, etc. since the beginning of the observation session). This process allowed us to assess whether birth attendants performed HH between patients. Observers entered this context-related information at the beginning of the observation session and updated it only if the situation changed.

Many of the recorded actions required further details to be entered. For example, when a delivery was observed, the observer also recorded
whether the delivery occurred rapidly (within 5 minutes of the woman walking into the labor room), whether there were complications, whether the observer birth attendant had an assistant, and whether a premade delivery kit was used. The observers collected contextual information and details of certain actions because we intended to use these data as potential determinants of HH in the analysis. The determinants collected and associated with HH are described in detail in Gon et al.24

### TOOL IMPLEMENTATION

This section characterizes how we used the tool to collect data and provides considerations for using it in future studies. We used the guidance provided by Zheng et al.25 for reporting time-and-motion studies, and include their full STAMP checklist information in Supplement 2.

### Sample Size Calculations

The data collection timeframe was based on the expected number of deliveries in the targeted facilities. We estimated the latter, using the formula for estimating a proportion from a cross-sectional survey with \( \alpha=0.05 \) and 80% power. We used a design effect of 2 based on a survey by Rowe et al.26 To estimate a hand rubbing/washing compliance of 10% with an absolute precision of ±3%, we needed 768 HH opportunities. We estimated the length of observation needed to collect this number, and in practice these data were collected during 336 observation sessions ranging from 13 minutes to 6 hours 45 minutes, with a median time of 1 hour and 41 minutes.8 As described in Gon et al.,8 we collected information on 781 HH opportunities before aseptic procedures (before aseptic procedures is 1 of the 5 types of HH opportunity prescribed by WHO).
Planning and Logistics of Data Collection

To obtain representative data on deliveries across all shifts (morning, afternoon, and night); 3 observers, 1 per shift, conducted observations that covered 24 hour a day. They observed for a total of 130 hours in the morning, 153 hours in the afternoon, and 205 hours in the night. Each observer had their own tablet for data collection. Each facility was visited for a mode of 6 consecutive days (range: 5–14 days) between September 17 and December 31, 2016. The order in which we visited the facilities was based on logistics. We arranged for additional days of observation in 1 facility with a high volume of staff to allow all staff to be observed and in 3 facilities with low volume of deliveries to capture a sufficient number of procedures.

We consulted the ward rosters to allocate individual attendants to the observers. Each attendant had a unique identifier that the observer had to record in WOMBAT when observing them. Observers were allocated to shifts based on the following principles: (1) the same observer should observe the same attendant so the attendant becomes accustomed to the same person being on the ward; (2) the initial attendant-observer pairs at each facility were assigned at random (unless specific concerns were raised; e.g., some flexibility on choice of types of shifts was allowed to accommodate observers’ needs); and (3) observation days should ideally be planned during changes in shift pattern to allow observation of the same attendants working on different types of shifts. The need to observe the same attendant across different types of shifts using the same observer increased the fieldwork duration and therefore had to be counterbalanced by the need to remain within our budget.

The Observers

Observers were all trained nurse-midwives working in managerial roles. Two of them worked in the study facilities but not in the labor wards. The third observer worked in district-level management. Their previous knowledge and understanding of the labor process were vital to ensuring quality during data collection and ultimately the project’s success.

Study Participants

All birth attendants present during the observation period who were involved in the childbirth procedures outlined in Table 1 were eligible for observation. We observed a total of 104 birth attendants across the 10 facilities and between 4 and 15 birth attendants in each facility. Each attendant was observed for 1–9 observation sessions. In each observation session, only 1 attendant was observed, but that attendant could be caring for multiple women and carrying out many procedures. Attendants in our study were all women, 90% were professionally trained, and 10% were health orderlies/nonprofessionals.

The attendants’ responsibilities were usually allocated during the shift itself. We encouraged observers to listen at staff meetings to learn which attendant was most likely to perform the childbirth procedures outlined in Table 1 to decide whom to observe. Observers were instructed to observe each allocated birth attendant roughly equally in each facility.

How to Observe

We trained the observers to enter only 1 action at a time to facilitate the data input process. We were specifically interested in the attendants’ actions, the sequence of these actions, and the length of time between them. We were specifically interested in the attendants’ actions, the sequence of these actions, and the length of time between them.

### TABLE 1. Relevant Hand Actions During Labor and Delivery Included in the HANDS at Birth Tool for Observation of Birth Attendants

<table>
<thead>
<tr>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measuring vital signs</td>
</tr>
<tr>
<td>Wiping the vagina</td>
</tr>
<tr>
<td>Vaginal examination</td>
</tr>
<tr>
<td>Artificial rupture of membranes</td>
</tr>
<tr>
<td>Episiotomy</td>
</tr>
<tr>
<td>Catching the baby (delivery)</td>
</tr>
<tr>
<td>Cord cutting and clamping</td>
</tr>
<tr>
<td>Cord traction</td>
</tr>
<tr>
<td>Postdelivery examination of the vagina</td>
</tr>
<tr>
<td>Wiping the baby clean after birth</td>
</tr>
<tr>
<td>Supporting breastfeeding</td>
</tr>
<tr>
<td>Manual removal of placenta</td>
</tr>
<tr>
<td>Suturing</td>
</tr>
<tr>
<td>Suctioning baby’s nose/mouth</td>
</tr>
<tr>
<td>Using bag and mask on the baby</td>
</tr>
<tr>
<td>Catheter insertion or removal</td>
</tr>
<tr>
<td>Insertion or removal of IV lines</td>
</tr>
<tr>
<td>Adjusting IV fluids or changing IV bag</td>
</tr>
</tbody>
</table>

Abbreviation: IV, intravenous.
each action per se. An action was selected and entered immediately. We do not have details on when the action ended, but since the actions were mutually exclusive, it was clear when one action replaced another.

**When to Observe**

As described above, a relevant patient-attendant interaction (Table 1) triggered the start of data entry; observers were expected to be continuously present in the ward due to the unpredictable nature of birth. Observers were encouraged to take breaks when no women were in labor or when women were in very early stages of labor and to remain where they could see if an emergency admission occurred to avoid missing delivery events. We also encouraged breaks if the observer’s concentration level was low.

We instructed observers to end a session when a major procedure ended and no further patient activities were in sight, when the observer wanted to take a break, when there was the opportunity to start observing another birth attendant, or when the birth attendants left the room to perform duties elsewhere.

**Where to Observe**

Observers would usually sit in the labor room. If no deliveries were happening, we asked observers to observe vaginal examinations in other rooms, such as the antenatal ward or examination room.

**Consent and Study Aim Concealment**

Written consent was gathered from women in the antenatal ward before observation; alternatively, women were asked for verbal consent once in the labor ward, and follow-up for written consent occurred in the postnatal ward before discharge or before delivery in the antenatal ward.² Women were told that no demographic information was collected on them and recorded observations were exclusively regarding birth attendants’ behavior. Permission to observe the attendants was obtained by the Ministry of Health and verbal consent was obtained by the observers when they first visited the facility.²

Attendants were told the observation was about the quality of care at birth, not on HH specifically, to conceal the study’s focus and reduce the Hawthorne effect. In all but the facility in Zanzibar where piloting took place, the focus of the study (HH practices) was likely to have been well concealed from the birth attendants being observed. The pilot facility in Zanzibar had the highest compliance with hand rubbing/washing before aseptic procedures. Compliance was 10% higher than the second-best facility and 7 times higher than the worst one. For ethical reasons, observers were trained to notify health workers and the field manager if they observed a potentially harmful condition or practice.

**Quality of Data Collection**

To ensure quality of data collection, we held regular meetings with collectors by telephone and onsite, communicated via a WhatsApp group, held Skype calls at the end of observations in each facility, and monitored the data uploaded monthly. These communication channels enabled rapid feedback, answers to questions, and maintenance of morale during long periods of observation. Drivers ensured observers arrived at sites on time. Finally, we are confident that the data were unlikely to have been manufactured because manufacturing time-stamped data would require as much time as conducting and recording actual observations.

**Software and Information Technology Costs**

The cost of the software and hardware also needs to be considered especially for deployment in low- and middle-income countries. WOMBAT 3.0 is available from the Apple Store (https://apps.apple.com/us/app/wombat-3-0/id1445107457). Data hosting is available at a cost of US$2,500 for a 2-year period, which allows the use of the software for multiple projects and data collectors. Free packages such as Open Data Kit could be used, but Open Data Kit is less user friendly for time-and-motion studies. In addition, we bought 3 tablets for approximately US$500.

---

**TOOL PERFORMANCE**

**Interobserver (Interrater) Agreement**

To report on interobserver agreement procedures and findings, we followed the recommendations by Lopetegui et al.²⁷ for time-and-motion studies and consulted the WOMBAT guidelines.¹⁸,²⁸ While piloting the tool, the trainer conducted 2 hours of simultaneous observation between the trainer (GG) and each of the observers. We then verified the extent of agreement between GG and each of the 3 observers on the basis of 28, 29, and 36 opportunities for hand washing/rubbing, glove wearing, and touch events, respectively. The observations were based on a total of 11 vaginal
examinations and 5 deliveries. The exercise was also used to provide feedback to the observers.

During the first month of data collection, we also assessed interobserver agreement, whereby a pair of observers was allocated to 2 of the same shifts in the busiest facility and asked to observe the same attendants. Observers were asked to perform this independently, avoiding communication or looking at each other’s tablet, but we could not ensure they were blinded, which meant that they probably knew we were going to check the data and hence some form of communication might still have occurred. Two pairs carried out this exercise for 1 morning and 1 afternoon shift each, the other pair for 2 night shifts. Two pairs observed 3 birth attendants, and the third pair observed 4.

We calculated kappa statistics based on either 49 or 50 hand rubbing/washing, hand recontamination, or glove behaviors per pair of observers. Observations were based on a total of 9 vaginal examinations and 11 deliveries. Through visual inspection of the data, we ensured that the behaviors compared were the same between observers by checking the reported time and sequence of actions. The kappa statistic calculated for pairs of observers was good for 2 of the 3 pairs at 0.93 and 0.90, but it was below the optimal level of 0.85 for 1 of the pairs at 0.73. In addition, we are also confident that discrepancies between observers was minimal because our results showed that hand rubbing/washing compliance before aseptic procedures did not vary substantially by observer, as described in Gon et al.8

Convergent Validity
We assessed the degree to which 2 measures of constructs that theoretically should be related were in fact related (convergent validity) by showing whether hand rubbing/washing before aseptic procedures compliance varied in the expected direction by contextual characteristics. Using the methods described in Gon et al.,24 we descriptively showed that higher compliance was present when the necessary equipment (water and soap or gel) was available, when fewer women were attended in the same observation session (i.e., a lower workload was expected to be associated with better HH), and when attendants had received HH refresher training in the previous year (Table 2).

### IMPLICATIONS FOR DATA ANALYSIS AND INTERPRETATION
In Supplement 3, we describe data cleaning, analysis, and interpretation issues that needed to be considered, noting in particular, that some data

<table>
<thead>
<tr>
<th>TABLE 2. Hand Rubbing/Washing Compliance Before Aseptic Procedures Among Birth Attendants in Health Facilities in Zanzibar, Tanzania</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed Opportunities/Indications for Hand Hygiene, per Guidelinesa</strong></td>
</tr>
<tr>
<td><strong>n (%)</strong>, <strong>N=779</strong></td>
</tr>
<tr>
<td><strong>Necessary hand hygiene equipment (water and soap or gel)</strong></td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>Missing</td>
</tr>
<tr>
<td>Inconsistent information</td>
</tr>
<tr>
<td><strong>Maximum number of women attended in an observation session</strong></td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>Missing</td>
</tr>
<tr>
<td><strong>Hand hygiene refresher training in the past 12 months</strong></td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>Yes</td>
</tr>
</tbody>
</table>

a Number of times when hand hygiene was meant to be performed per guidelines.
items relied on observer subjectivity (e.g., duration of hand washing) and some variables (e.g., variables describing the context) required more stringent training than others.

**Data Structure**
A strength of WOMBAT is that when each action is recorded, the time of that action is automatically logged. Our final dataset was a list of 7,893 time-ordered entries. These data were coded to derive HH opportunities and to calculate compliance. First, each HH opportunity needed to be identified within each observation session, which is further explained below. Second, for HH opportunities before aseptic procedures or touching the patient, the sequence of actions preceding the opportunity needed to be examined for hand rubbing/washing actions, glove use and actions that may lead to a new HH opportunity. Whereas, for HH opportunities after exposure to body fluids or touching the patient or the patient’s surrounding, the actions following the opportunity needed to be examined. We used STATA to analyze these data.

**Time Stamps**
We used WOMBAT’s time stamp information in 2 ways. First, we checked the plausibility of certain actions being linked; for example, a hand rubbing/washing action could not be linked to a procedure conducted 10 hours before or after it. Second, we calculated the length of time between hand rubbing/washing and the HH opportunity to determine whether time would predict the likelihood of hand recontamination occurring.

**A Priori Definitions Required**
To estimate HH compliance, we operationalized definitions for the systematic flow of patient contacts allowed within a given HH opportunity and the patient zone. By a systematic flow, which we called a “delivery flow,”8 we referred to the procedures or actions of interest that defined the start of a new HH opportunity, as well as the sequence of these procedures, which occurred without a break and were considered as 1 opportunity for HH.8 For example, in a given delivery flow, a vaginal examination could be followed by the delivery of the baby, but not by touching a patient’s shoulder. During a delivery flow, a birth attendant could undertake hand actions within the patient zone without the need for a new HH opportunity to arise.

In this study, we defined patient zone as encompassing a woman’s perineal area and thighs, any clean or sterile equipment being used, and the newborn as it was caught and wiped. A break in the delivery flow, indicating a new HH opportunity, arose if an activity occurred that was outside the patient zone, such as inserting an intravenous line, touching the patient beyond the zone, or leaving the room.8

Details on the definitions used in our study are reported in Gon et al.8 Potentially, a separate software could be programmed to automatically analyze this type of data in the future, allowing for definitions to be applied from the outset.

**Context-Specific Adaptations**
To classify which surfaces we should include in the patient zone, we used previous formative research on the microbiological load of the labor surfaces in Zanzibar, as well as unstructured observation of labor wards conducted within the HANDS project. For example, we excluded the delivery bed and trolley from the patient zone because previous work found that these surfaces were often contaminated with potential pathogens.8 Other important information to consider include the details of the cloth or plastic sheet used under the woman’s body during birth, the cleaning routines of the wards, the type of water available, the delivery equipment preparation, and the local HH guidelines against which to measure hand washing/rubbing duration and technique. It is not clear that all projects will have the capacity to gather this level of contextual information; however, capturing the real workflows in this context was our aim.

Ideally, all definitions should be clear at the start of a project, but during data collection, the project may accrue context-specific information on the surfaces or the attendants’ workflows, which should be used to update the definitions. To illustrate this, we present the number of HH opportunities and hand rubbing/washing compliance results for 4 different patient zone definitions (Supplement 4).

**DISCUSSION**
We developed the HANDS at Birth tool to capture the complex HH and glove behaviors of birth attendants, based on state-of-the-art methods: a time-and-motion study using a computerized system (WOMBAT). This approach has been rarely used to measure HH or to conduct research in low-resource settings.10,19,20,30,31 Our time-and-motion study enabled us to accomplish the following, which would have not been possible with the
WHO HH Observation Form: (1) to look at whether birth attendants comply with the complete sequence of behaviors prescribed by the WHO guidelines,\(^8\) (2) to look at each behavior individually, and (3) to look at different behavior sequences.\(^8\) Additionally, our method reduced the risk of observer bias because data collection was coded as a series of individual actions rather than relying on observer judgment that a new HH opportunity had occurred; hence, opportunities were identified retrospectively in a standardized way.\(^8\) Indeed, hand rubbing/washing compliance was similar between observers in our study, as reported in Gon et al.\(^8\) Beyond HH, the HANDS at Birth tool allowed investigation of other behavior sequences and workflows.

We are aware of 1 other study that used time-and-motion methods to report HH of health care workers in the context of an intensive care unit in the United States.\(^31\) That study’s aims differed from ours including determining the number of contacts between patients and health care workers, as well as how long they take, and estimating HH compliance specifically before entering a room and after exiting a room. That study did not detail information on the tool format or content. In comparison, the HANDS at Birth tool allows for a more exhaustive list of actions to be recorded, including those beyond patient-attendant interactions; it also allows looking at all HH opportunities, not just those related to entering or exiting the room.

This tool has the potential to be adapted to examine HH in other types of wards. We think this detailed examination of HH, including recontamination, is particularly important in wards facing unpredictable volumes of patients or unpredictable patient complications. Examples include emergency departments, operating wards, or isolation wards during epidemics, such as the current isolation wards for COVID-19 patients. In particular during the COVID-19 pandemic, this tool could lend itself to examining the key relationship between hands and surfaces and the fundamental issue of pathogen cross-contamination between them.\(^34,35\)

**Limitations**

Because we were interested in individual determinants of HH behavior, we observed only 1 birth attendant at any 1 time; whereas, the WHO HH Observation Form audit tool is designed to observe multiple health care workers simultaneously, which allows collection of more HH opportunities in the same observation session. Important, the HANDS at Birth tool is not intended to substitute for the WHO HH Observation Form; the 2 tools serve very different purposes, with the former being aimed at research and the latter at infection prevention practitioners. Another limitation of our tool, and how we used it, is that it requires data cleaning and data management. For example, even though misclassification was minimal, some actions were recorded by mistake at the same time. In addition, a couple of variables relied on observer subjectivity—for example whether a delivery happened very fast after the woman’s admission in the labor room. The structure of the data implies that data management is needed to create HH opportunities and HH compliance results.

**CONCLUSION**

In conclusion, we report the process of developing a research tool to capture the complexity of HH and glove behavior during labor and delivery, including the tool elements, field implementation, tool performance, and implications for analysis. We used a computerized system that was feasible to use in low-resource facilities. Advantages of this tool include simpler training, less observer bias in assessing HH compliance (compared with the WHO HH Observation Form), and the ability to monitor multiple behaviors. The data it produced also showed good reliability and convergent validity. Future studies should explore the use of this research tool in labor wards in other contexts, as well as in other types of wards.

**Acknowledgments:** We thank the Ministry of Health of Zanzibar for their participation and engagement in the study. A special thanks to Rukaiya M Said, Mwanatafima Ali Mohammed, Biju Mubwua Abdallah, and Asya Hati Vulai who collected all the data. We also thank Marina Daniele for participating in the consultation exercise aimed at refining the definition of hand hygiene opportunity.

**Funding:** The project was funded by the United Kingdom Research & Innovation Medical Research Council (MRC) Public Health Intervention Development scheme (award number MR/N015975/1). This award is jointly funded by the MRC and the United Kingdom Department for International Development (DFID) under the MRC/DFID Concordat agreement and is also part of the European & Developing Countries Clinical Trials Partnership program supported by the European Union. The Soapbox Collaborative also contributed by funding staff involved in this project.

**Competing interests:** None declared.

**Data Availability:** The dataset generated during the current study is available at: https://doi.org/10.17037/DATA.00000778.

**REFERENCES**


17. Westbrook JI, Ampt A, Design, application and testing of the Work

14. Yawson AE, Hesse AAJ. Hand hygiene practices and resources in a
CrossRef. Medline


Implementation of a Pediatric Early Warning Score to Improve Communication and Nursing Empowerment in a Rural District Hospital in Rwanda

Shela Sridhar, a Alexis Schmid, a Francois Biziyaremye, b Samantha Hodge, a Ngamika Patient, c Kim Wilson a

Key Findings

- Nursing staff felt more empowered to communicate clinical findings to the physician team following the Pediatric Early Warning Score for Resource-Limited Settings tool training and implementation process.
- The process of implementing the tool triggered more calls from nursing staff to the physician teams to initiate early intervention.

Key Implications

- The Pediatric Early Warning Score for Resource-Limited Settings tool has the potential to improve competency and confidence of nurses in their triage capabilities. Although traditionally implemented in tertiary care centers, program managers should consider implementing this tool at the district hospital (secondary) level as well.
- Physicians and nurses both play crucial roles in triaging systems. Therefore, program managers should consider engaging both groups with the tool before and during implementation.

ABSTRACT

Background: Pediatric early warning (PEW) scores represent a “track-and-trigger system” that identifies clinical deterioration in a patient’s condition in the hours preceding a sentinel event. Before implementation, nurses reported feeling unprepared to identify and advocate for acutely ill patients owing to a lack of skills, vocabulary, and agency. We implemented a Pediatric Early Warning Score for Resource-Limited Settings (PEWS-RL) with nurses in a rural district hospital in Rwanda. Although PEW scores can improve clinical outcomes, empowering nurses in resource-limited settings to discuss patient acuity with physicians is a critical first step. Our primary aims were to train nurses to obtain more accurate vital signs and assess their importance as early warning signs of clinical deterioration and use PEWS scores to improve communication between nurses and physicians.

Implementation: The PEWS-RL tool implementation began with a training program that was created through discussions with nurses, physicians, and the medical director of the hospital. The program included lectures and application of learned skills through direct clinical mentorship of nurses, as well as training of physicians regarding PEWS-RL as a communication tool.

Evaluation: The PEWS-RL protocol was evaluated based on pre- and post-tests to assess improvement in nurses’ knowledge and skill, as well as skills assessments of accurate recognition of clinical deterioration. All 6 nurses passed skill testing with >80% accuracy. Nurses’ feelings of empowerment to advocate for patients and to escalate care were assessed through pre- and post-training interviews. Nurses described increased confidence in calling for physician support.

Discussion: Implementation of PEW scores increased nurses’ technical skills and feelings of confidence and empowerment; however, the low-resource setting presented major challenges. Barriers to sustainable implementation include the rapid ward staff turnover as well as limited physician buy-in. Nevertheless, the PEWS-RL tool has the potential to empower nurses and improve patient outcomes if fully embraced by staff.

BACKGROUND

Responding to Increased Child Mortality Rate

Kirehe District Hospital (KDH) is a public hospital in rural Rwanda, supported by a partnership with the nongovernmental organization (NGO) Inshuti Mu
Buzima, a local organization of Partners in Health. KDH serves a catchment area of approximately 340,000 people; or 13% of Rwanda’s population, including a large refugee settlement with 57,000 residents.1 About 50% of the population in Kirehe is under the age of 17. KDH has a busy general pediatric ward, with 60–120 pediatric admissions per month of children aged 1 month to 15 years old. The staffing model for the pediatric ward includes 1 or 2 nurses caring for 10 to 30 pediatric patients, supported by a general physician covering the pediatric ward as well as a 40-bed neonatal ward. The total medical staff comprised 6 pediatric nurses and 10 physicians. In addition, a U.S.-trained pediatrician affiliated with Partners in Health works with the hospital to conduct on-site clinical capacity building for several months each year.

In 2018, hospital staff noted rising mortality rates in the pediatric ward. A chart review between May and October 2018 indicated that for some months the child mortality rate was as high as 6% and the average for the 6 months about 3%. The majority of deaths were associated with sepsis or pneumonia resulting in respiratory failure, often as a consequence of inadequate recognition of altered mental status and respiratory fatigue. Death occurred an average of 7 days after admission, with a minimum of <24 hours and maximum of 27 days after admission. Potential etiologies for death after 7 days of admission may have been iatrogenic (management of patients), but we suspected there was also poor recognition of warning symptoms and clinical progression of disease that may have been missed in the days leading to mortality later in the hospital course.

A variety of factors can contribute to clinical deterioration or high rates of mortality in children with sepsis or pneumonia. However, the incidence of such events can be significantly lower in similar settings where early recognition of deterioration and prompt initiation of treatment or early transfer to higher levels of care have been initiated.2

More than 95% of pneumonia-related deaths occur in low- and middle-income countries. However, little data exists on quality-of-care indicators and practices around pneumonia care in these contexts.3 This paucity of data suggests a gap in assessment of quality of care in pediatric populations.

Understanding Hospital Care Delivery Factors
To better understand hospital-based care delivery factors that could be contributing to high pediatric mortality, we conducted key informant interviews with the 6 pediatric nurses and 4 physicians. After observing multiple deaths on the pediatric ward, the Partners in Health-affiliated pediatrician hypothesized that a driver of the mortality rate was inadequate communication between nurses and physicians. This was observed at 2 separate Partners in Health supported hospitals. However, because KDH was the busier facility at the time, it was chosen for the intervention.

We conducted informal interviews with 4 physicians to gain a background understanding of physicians’ perceptions regarding nursing competencies. The 4 physicians interviewed were those who spent the greatest amount of time on the pediatric ward. We asked the 6 pediatric nurses questions regarding their comfort level with triaging and communication processes. The interviews also included open-ended questions to draw out additional themes. We asked the interview questions in English and used an interpreter who spoke Kinyarwanda and English to translate. Interviews with nurses and physicians lasted no more than 30 minutes. Common themes that emerged for both nurses and physicians included constraints on time and human resources, which compromised the clinicians’ ability to appropriately prioritize patients and complete tasks. Nurses also cited gaps in knowledge and skills in identifying and subsequently reporting the status of critically ill patients as creating a barrier to timely care. Nursing leadership highlighted that nurses felt disempowered to advocate for deteriorating patients. They reported lacking a common language around assessment of critical illness with physicians and therefore feeling unprepared to highlight the acuity of a patient’s condition and effectively advocate for them. Other nurses reported that their concerns were sometimes dismissed by physicians who would respond by saying:

*Pediatric vitals are different.*

General physicians cited concerns with the accuracy of vital signs reported by nurses as a key barrier in assessing the severity of pediatric illness. Multiple physicians indicated that:

*If I want to believe the vitals, I take them myself.*

Using Pediatric Early Warning Score for Triage
Although time constraints and human resource allocation are subject to financial constraints, clinical processes such as accurate triage and communication are modifiable factors with minimal financial burdens on low-resource hospital systems. Pediatric early warning (PEW) scores represent a
triage “track-and-trigger system” that can accurately identify up to 85% of children who will experience clinical deterioration, such as cardiac arrest or severe respiratory compromise, sometimes as early as 11 hours before the sentinel event. The scores are a mechanism that can be used to modify triage systems and standardize communication regarding acutely ill children. PEW scores have primarily been used in high-income countries, but they were recently adapted for use in resource-limited settings (Pediatric Early Warning Score in Resource-Limited Settings [PEWS-RL]) and validated in a tertiary care setting in Rwanda. The PEWS-RL uses basic clinical assessments including respiratory rate, respiratory distress, heart rate, temperature, blood pressure, oxygen use, and mental status. It demonstrated a 92% sensitivity and an 87% specificity in identifying children at risk of clinical deterioration.

To improve early recognition and communication of clinical deterioration in pediatric patients by nursing staff, we aimed to implement a standardized triage system including a standardized clinical assessment for patients at risk for clinical deterioration in our inpatient pediatric ward. We hypothesized that this intervention would improve nurses’ ability to accurately identify critically ill patients, improve communication about critical patients by creating a common language with physicians, and prompt a timely physician response to evaluate and initiate the appropriate medical management for a child whose condition is deteriorating. In this report, we describe our process for implementation of the PEWS-RL at the district hospital level, including areas of success, challenges, and lessons learned.

IMPLEMENTING PEW SCORES

We reviewed several versions of the PEW triage tools collaboratively with staff and leadership at KDH, including the medical director, clinical director, and the primary general practitioner who rounded on the pediatric wards. PEW systems include 2 components: a score calculated using vital signs at prescribed intervals during a child’s hospitalization and a response system, which may be as simple as contacting a physician, that is activated if a specific score threshold on the tool has been reached. Early warning scores commonly evaluate and score vital signs as well as clinical exam assessments, such as level of consciousness, capillary refill, or work of breathing. No general consensus exists regarding which components are essential, the frequency with which they should be recorded, or the thresholds and scoring mechanism that indicate clinical concern. Few versions have been evaluated in resource-limited settings where staffing ratios and the level of nurses’ training differ. After an analysis of PEW tools and initial conversations with staff, we determined that it would be best to focus on objective data (basic vital signs alone) without including clinical assessment. Given that the PEWS-RL met these criteria and had previously been validated in Rwanda in a tertiary hospital setting, we decided to implement the same version in our hospital for consistency and potential nationwide scalability in the future.

The PEWS-RL tool (Figure) is purposefully composed solely of vital signs that are attainable with minimal equipment or assessment ability. This approach was taken because clinical assessments, such as blood pressure and respiratory effort, are often not examined due to a lack of trained personnel and availability of pediatric-sized equipment. This tool was utilized across the pediatric age range (1 month to 15 years). Our PEW score included respiratory rate, heart rate, temperature, and mental status; each was scored at 1 point. Physician notification was triggered at a score of 3 on admission or an increase of 3 points on subsequent assessments. Blood pressure was initially included in the assessment of the PEWS-RL; however, based on discussions with the research team at the University Teaching Hospital of Kigali, the sensitivity and specificity of the tool did not notably change when blood pressure was removed. Therefore, we did not include it in our score.

Training Approach

The initial implementation of PEW scores started with a training program for 6 nurses and 10 physicians over a 2-week period in November 2018. A visiting U.S.-based pediatric nurse specialist provided 1–2 hours of on-site didactic training per day to the pediatric ward nursing staff, focusing on the clinical importance and implications of abnormal vital sign values. Although many (but not all) physicians and nurses are exposed to emergency triage and assessment training in school, ongoing mentorship or recertification is uncommon. The remainder of the day focused on application of learned skills and direct clinical mentorship through “real-time” patient assessments and active feedback at the bedside. Nurse training included bedside mentoring during morning rounds and didactic sessions each afternoon with continued bedside teaching through the day. The training was
FIGURE. Pediatric Early Warning Score for Resource-Limited Settings Tool Used at the University Teaching Hospital of Kigali, Rwanda, and Kirehe District Hospital

Abbreviation: BP, blood pressure.
incorporated into the nursing workflow to the extent possible to be minimally disruptive in an already understaffed environment.

Along with clinical and didactic training, nurses were provided stethoscopes so they could manually check heart rates and blood pressures as a secondary validation of the cardiorespiratory monitor used on the unit. They were also provided individual pulse oximeters to be used as a secondary check on the existing monitor. Equipment was distributed at the start of the training and used throughout the 2-week course.

General physicians were given 2 lectures dedicated to understanding the PEW score and response system by the NGO pediatrician. Lectures for nurses focused on assessment and reporting of the score, and lectures for physicians focused on responses to different scores and critical thinking around common case scenarios. In addition, the pediatrician rounded with the clinical team each morning for the 3 months following the training program and provided ongoing low-dose, high-frequency mentorship for both rounding physicians and nurses in using and interpreting the PEWS-RL. PEW scores for patients were reviewed each morning to assess for completion and to facilitate discussion of any challenges encountered during the implementation. Through this process, PEW score documentation was integrated into ward rounds and the existing workflow for both nurses and physicians.

Assessment of Outcomes
Nursing and Physician Knowledge, Skills, and Clinical Practices

We conducted an evaluation of the impact of PEWS-RL implementation at KDH on nursing and physician knowledge, skills, and clinical practices using interview data and process measures. Our measures included changes in nursing knowledge and skill in accurate recognition of clinical deterioration, changes in nursing physician communication before sentinel events, and nursing confidence levels in the communication of clinical findings to physician staff. The PEWS-RL triage system includes both the risk score obtained by the nursing staff and the responsiveness of the physician team. Our assessments focused on the primary objective of nursing competency and communication rather than the response component of triage systems.

Nursing knowledge and skills in accurate recognition of clinical deterioration was evaluated using pre- and post-training written tests and clinical skills assessments. Skill competency of KDH pediatric nursing staff was evaluated by the pediatric nurse specialist using a standardized checklist immediately after the 2-week training program. The objective exam focused on their ability to obtain manual vital signs and calculate a PEW score. The clinical competency form used during the assessment is outlined in the Supplement.

A numeric score as well as written feedback was provided. All 6 pediatric nurses independently passed skill testing with >80% accuracy. In addition to the clinical skills assessment, a written exam was given on the first and last day of training to assess the clinical knowledge necessary to adequately utilize the PEWS-RL as intended for screening and response activation. The average pretest score for nurses was 66% with a range of 53%–80%. The average post-test score was 81% with a range of 67%–100%.

Nursing Communication with Physicians

We conducted a qualitative assessment of the impact of our training on nursing communication with physicians, focusing on the nurses’ level of empowerment in patient advocacy and escalation of care around sentinel events. Pre- and post-training structured interviews of the nursing staff were conducted by an NGO nurse mentor individually and confidentially and in their primary language to promote more open communication. In addition, nurses completed a written survey of their communication practices and comfort level in escalating care before and after completing the 2-week training program. The survey included Likert-scale questions with declarative statements such as “I feel comfortable asking physicians questions” as well as free-form answers to questions around communication such as “How do you feel communication between nurses and physicians affects patient care?” Before training, 2 of 6 nurses felt as though their clinical assessments were often dismissed by physicians. Post-training, nurses stated feeling more confident in their ability to advocate for patients. One nurse articulated:

Now I have a tool to back me up when I call the doctor.

Nurses described increased confidence calling for physician support:

[We] have something to say.

Changes in nursing clinical communication practices were also evaluated by retrospective chart review. In the 6 months before the implementation of our PEWS protocol, only 8 of 30 patients (27%)
Implementation of a Pediatric Early Warning Score in Rural Rwanda

who had been transferred or died had a recorded call to a physician in the 11 hours preceding the sentinel event (transfer to a tertiary care center, resuscitation by KDH staff, or death). However, in the 2 months immediately following the intervention, we found that the physician was called 63% of the time (7/11 patients) before a sentinel event. In the 6 months before the intervention, there was a 3% mortality rate and a 3% transfer rate for all patients in the hospital. In the 2 months after our intervention, the mortality rate was not significantly different, but the transfer rate had increased by 11%. A chart review also demonstrated an increase in physician response frequency; however, physician response times will need to be evaluated over time in subsequent iterations. Although the data collected from this training program are not adequately powered to make definitive conclusions regarding the effect of PEW scores on clinical outcomes or physician response times, the initial inferences are promising and merit further investigation.

**DISCUSSION**

Implementation of our PEW score created an opportunity for vital nursing education on high-quality assessment of vital signs and a deeper clinical understanding of underlying pediatric physiology. The PEWS-RL implementation also empowered nurses, provided them with a model, and mentored them on the tools to communicate their assessments to physicians. The public hospital-NGO partnership provided the opportunity for nurses and physicians to observe communication between an external nurse and physician, which modeled important skills in open multidisciplinary communication and physician trust in nursing assessment skills. Given the immediate increase we saw in recorded calls to physicians being documented and postintervention interviews with nursing staff, the PEWS-RL tool may be effective even on a busy ward with limited staff and resources.

**Challenges and Lessons Learned**

Although nursing knowledge and skills demonstrably improved and nursing staff reported feeling empowered, we encountered several challenges during the implementation of our protocol. Our most significant challenge was in motivating physician engagement. We provided lectures for physicians over the course of the training, but our interactions and reports from nursing staff suggested that the physicians were less interested than the nursing staff in using the PEWS-RL. The reasons for this finding are likely multifactorial. The primary focus of our efforts in training was toward the nursing staff, and we included only 2 training sessions and no formal mentorship process for physicians who did not round on the pediatric ward during the implementation phase. Additionally, of the 16 physicians on staff at the time of the intervention, only 4 were able to spend time on the pediatric wards during the initial implementation of the protocol. Finally, as our focus was on nursing empowerment, we failed to involve the physicians in the planning process of protocol development and implementation, likely leading to inadequate physician understanding and involvement in the protocol.

The critical lesson learned was the importance of engaging physicians and nurses together. We were able to implement the first steps in a triage system, 8 that is, nursing recognition and empowerment to communicate clinical findings. However, changes to outcomes and mortality will require true physician engagement and understanding of the PEWS tool, its implications, and how to respond to nursing concerns. As noted in the PEW score literature, to be effective in reducing morbidity and mortality, the tool needs to be implemented within a system that is able to respond to the needs of the child: specifically, a provider or a team that has the ability to not only accurately assess the patient and recognize anomalies, but to also implement appropriate clinical interventions. 7,9,10 Our PEWS intervention focused on addressing the first steps of recognition and empowerment. However, we did not address the subsequent step, which requires “the assistance is readily available and appropriately skilled...” 8 Physician engagement can be further accomplished by ensuring leadership and ownership of the tool implementation by the physicians. Additionally, appointing a physician leader in pediatrics may be useful to create physician buy-in and organizational accountability.

Although we attempted to educate physicians and nurses side-by-side during medical rounds, we should have placed a greater focus on individual-physician coaching to address the aspect of skilled assistance in response to recognition of illness. Furthermore, before implementation, we did not adjust the physician schedule to ensure that all medical staff rotated through the ward during the training. This decision was made to minimize disruptions to work flow in the hospital as well as minimize administrative burden on the clinical director who managed the schedules. This reasoning also informed why we did not have...
additional meetings during the training with the entire staff to review and address problems that were noted, but rather managed them on an ad hoc, daily basis. However, in subsequent programs, the short-term disruption may be acceptable if longer-term clinical benefits can be derived.

Other barriers to sustainable implementation of this protocol included the rapid turnover of staff on the ward and a loss of equipment following the training. The chief of nursing was transferred during our training and 1 week later another member of the nursing staff left and was replaced by a new staffer. Next steps to mitigate this limitation include creating an on-boarding system for pediatric nurses as well as new physician staff. Additionally, to minimize the removal of stethoscopes and oximeters from the ward, they should be tagged with large, bulky labels or affixed to a mobile cart, or the workflow should be changed to include retrieval of locked equipment at the start of each nursing shift. Although we are encouraged by the feedback from our nursing staff, another limitation was our status as an outside organization. Survey answers may have been biased due to cultural tendencies to avoid criticizing the system or a desire to provide positive answers.

Limitations
In addition to the procedural challenges we experienced, important limitations in the assessment and interpretation of the outcomes should be considered. Our assessment focused on nursing competencies, with minimal evaluation of physician responses. This approach prevented us from measuring clinical outcomes and timeliness of responses, which are critical to any effective triage system. Additionally, the program was conducted at one hospital with a small number of nurses. To achieve statistical significance for nursing competencies, this program would need to be conducted across multiple sites or over multiple iterations at KDH. Finally, although the PEW score has been validated in tertiary care centers where few interventions are available to inform the medical response, it has not yet been validated in district-level hospitals where fewer interventions are available to inform the medical response.

SUMMARY AND CONCLUSIONS
Early recognition and response to clinical deterioration are essential to improving outcomes in pediatric care in hospitalized patients in resource-limited settings but can be challenging. Initial barriers can include limitations in nursing knowledge and skills in pediatric triage and lack of nursing empowerment in escalating concerns for timely physician response. Training and implementation of the PEWS-RL resulted in demonstrable improvements of both technical skills and feelings of confidence and empowerment among the nursing staff. Challenges and next steps in quality improvement and implementation remain, including the need to address equipment availability and security and the implementation of approaches to improving physician training and buy-in.

However, it remains to be seen if the subsequent steps of the track-and-trigger system can be improved with increased physician involvement in the implementation process. This next step is crucial given that our triage system would ultimately be incomplete without an appropriate response and intervention system. The next steps are to design an adequately powered study across multiple district hospitals to evaluate the feasibility and effect of PEW scores in low-resource rural settings. These studies should start by focusing on delivery of care at the hospital level including physician–nursing communication, response times, and appropriateness of medical management to PEW triggers. Based on our implementation experience, it will be essential to conduct intensive nursing and physician training simultaneously with a dedicated review process. Nevertheless, based on our initial assessment, the implementation of PEW scores in a rural district hospital in sub-Saharan Africa has the potential to empower nurses and improve patient outcomes in low-resource settings if fully embraced by staff.

Acknowledgments: Thank you to Catherine Kirk and Inshuti Mu Buzima in Rwanda for their continued support and assistance with our training program. We thank the pediatric staff at Kirehe District Hospital for their willingness to learn and take an additional work despite their overwhelming existing patient care responsibilities. Thank you to the Global Health Department at Boston Children’s Hospital for financial and technical support throughout the process.

Competing interests: None declared.

REFERENCES


Peer Reviewed

Received: February 13, 2020; Accepted: August 21, 2020; First published online: October 15, 2020


© Sridhar 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-20-00075
Juntos: A Support Program for Families Impacted by Congenital Zika Syndrome in Brazil

Antony Duttine, Tracey Smythe, Miriam Ribeiro Calheiros de Sa, Silvia Ferrite, Maria Elisabeth Moreira, Hannah Kuper

Key Messages

- A community-based family group program for caregivers of children with congenital Zika syndrome (CZS) was developed based on an existing program for families of children with cerebral palsy and pilot tested in Brazil.
- Program managers developing group initiatives should consider fast-track learning approaches to adjust their intervention to make it more useful to participants.
- Clinicians and managers of Zika initiatives should consider that caregivers of children with CZS will likely benefit from the ability to engage and exchange with caregivers of children with other neurodevelopmental disabilities.
- Program managers delivering community interventions targeting caregivers should consider engaging expert mothers as group cofacilitators.
- Integrating emotional support activities into groups that address child development is important, and adds value.

ABSTRACT

Background: The 2015–2016 Zika virus outbreak in Brazil was unprecedented and resulted in the birth of more than 3,000 children with congenital Zika syndrome (CZS). These children experience multiple complex health conditions and have limited services to support them and their family’s needs.

Program Development and Piloting: An existing family support program for children with cerebral palsy (Getting to Know Cerebral Palsy) was adapted to the Zika context in Brazil through expert consultation. The program was pilot tested at 2 sites among 6 groups of caregivers (total of 48 families) from August 2017 to June 2018. Group observation and focus group discussions with facilitators and participants at the end of each session informed fast-track learning, which was used to tailor the program for future groups. Fast-track learning—adjusting the intervention in real time based on gathered feedback—was found to be a helpful process to inform and hone the program from its initial concept.

Program Description: The intervention, Juntos, is a facilitated participatory group program for caregivers of children who have CZS. The group sessions are cofacilitated by a parent of a child who has CZS and an allied health professional. The group meets for 10 sessions that last 4 hours. Each session includes an icebreaker, activities, and group discussions. Content covers practical information on caring for a child with a developmental disability including that caused by Zika. Psychosocial support forms an important component, and families are guided from the first week to define and develop their own communities of support. Six pilot groups were successfully run in Rio de Janeiro and Greater Salvador, Bahia. The groups gave positive feedback on acceptability and demand.

Conclusions: The program has the potential to be an important tool for community health and social support services in South America in response to Zika. The program can also be applied to children with neurodevelopmental disabilities other than those caused by the Zika virus, which could be important in ensuring families of children with CZS are less isolated.

BACKGROUND

The Zika outbreak of 2015–2016 in South America caught the international health community unaware. There had previously been no severe health consequences associated with the virus, despite Zika having been known since the 1940s.1,2 Zika has now been
proven to cause developmental impairments in children, collectively known as congenital Zika syndrome (CZS). This syndrome includes microcephaly as the most pronounced and documented symptom, which is linked with severe and multiple impairments. Evidence is emerging that Zika also causes an array of other cognitive and physical impairments that may not be immediately apparent at birth. Microcephaly is likely to be the tip of the iceberg in terms of affected children, as more mild or moderate impairments stemming from in utero Zika infection appear to be far more frequent. Brazil was the most affected country in the outbreak. As of March 2020, Brazil had 3,559 confirmed cases of CZS with an additional 2,871 cases under investigation (total 6,430 cases).

Although CZS and cerebral palsy are separate conditions, because they have similarities, programs designed for caregivers of children with cerebral palsy could provide a strong foundation to adapt a program for the Zika context in Brazil. One such program, Getting to Know Cerebral Palsy (GTKCP), was developed by the London School of Hygiene & Tropical Medicine (LSHTM) after a childhood disability survey showed that caregivers of children with cerebral palsy in Bangladesh had very little access to information or support regarding the best way to care for their child and that available services were extremely limited. GTKCP is a 10-session parent-support program held in the community that aims to improve parents’ knowledge and skills in caring for their child and improve the quality of life of parents and children with developmental disabilities. It is hard to estimate the exact reach of the program, but an online community of practice established in 2014 to support the rollout of GTKCP has 412 members across 72 countries who share knowledge and experiences. GTKCP focuses on parents of children aged 2 years and older; a new version, the Early Intervention Program (EIP), was developed for parents of children aged younger than 2 years. Program material is available from www.ubuntu-hub.org.

Needs Analysis
From April to August 2017, we conducted a needs analysis to assess the potential value of a community-based program, based on GTKCP, for caregivers of children with CZS in Brazil. The needs assessment involved: (1) tracking and comparing emerging literature on the clinical presentation of CZS with existing literature on cerebral palsy; (2) conducting a literature review on the needs of caregivers of children with CZS and cerebral palsy in middle-income contexts; (3) meeting with caregivers, specialists, and other local stakeholders in Brazil to identify key gaps, challenges, and needs; and (4) reviewing emerging data from a sister study measuring the social and economic impact of CZS on caregivers. A full description of the needs analysis is available.

We found that providing some services for children with complex multiple impairments at the community level could be crucial to address the unmet needs experienced by families of children with CZS in Brazil and may be more affordable than centralized services (which may be difficult or costly to access). Families of children with CZS, particularly those children with more severe impairments, did not have enough access to specialized health and rehabilitative services and informal support groups, and formalized support for caregivers was also limited. There was some concern raised by clinicians that children with mild to moderate impairments stemming from Zika infection were less likely to attend rehabilitation and that these caregivers were an important group to be targeted. Other researchers have also reported on the additional services required to fully address the care needs of children with CZS and their families.

Given the results of the needs analysis that identified the unmet support needs of parents in Brazil and the positive reception of the principle of GTKCP for Brazil among local stakeholders, researchers at the LSHTM who had been involved in GTKCP and EIP felt that adapting GTKCP and EIP for the Zika context and Brazilian culture could be potentially useful. Partnership for the project was established between the LSHTM and 2 Brazilian institutions: the Instituto Nacional de Saúde da Mulher, da Criança e do Adolescente Fernandes Figueira (IFF) in Rio de Janeiro, and the Universidade Federal da Bahia (UFBA) in Salvador.

This article describes the process of developing and piloting the intervention in Brazil, as well as the final program that was developed (Figure 1). We also reflect on lessons learned as key recommendations from this innovative program may be useful for other global health practitioners designing community-based family group interventions.
A protocol was established by the lead project researcher (AD) to measure feasibility of the pilot intervention.13

**Ethics Approval**

Ethical approval was obtained from the Instituto de Saúde Coletiva/UFBA Ethics Ref 2.369.348, IFF/FIOCRUZ RJ/MS Ethics Ref 2.183.547, and LSHTM Ethics Ref 13608. Informed consent was acquired from all participants.

**Initial Adaptation of Program**

To support the adaptation, advisory groups were established in Brazil and in the United Kingdom and included a range of specialists, as well as mothers of children with CZS.

The GTKCP and EIP curricula were reviewed by the lead project researcher (AD) with other LSHTM colleagues (TS, HK), Brazilian colleagues (SF, MS), the GTKCP and EIP teams, and other key identified experts (including specialists). The project lead is a physiotherapist with 15 years of programmatic experience, including in qualitative and participatory research and community-based rehabilitation in low- and middle-income countries. During a May 2017 workshop in London, the experts convened to discuss the preliminary findings of the needs assessment and to develop consensus on a first draft outline of the program, an initial timeline, constituency of the facilitators to lead the caregiver group sessions, and participant inclusion criteria.

The project group developed a theory of change to describe how the program relates to broader societal participation of children with developmental delays, including CZS, and the pathways that determine the extent to which this intervention may be successful. The theory of change describes what changes are needed and the assumptions underlying the achievement of these changes.18 Therefore, the theory of change linked outcomes with activities to explain how and why the desired change was expected to occur.

**FIGURE 1.** Timeline of Juntos Program Development for Caregivers of Children with Congenital Zika Syndrome, Brazil

<table>
<thead>
<tr>
<th>Needs Analysis</th>
<th>Pilot Phase 1: 2 groups</th>
<th>Pilot Phase 2: 4 groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literature review, interviews</td>
<td>Fast-track learning, program adjustment, data collection</td>
<td>Ongoing fast-track learning, Data collection</td>
</tr>
<tr>
<td>April 2017</td>
<td>May 2017</td>
<td>June 2017</td>
</tr>
<tr>
<td>• Adaptation Workshop</td>
<td>• Researcher and facilitator initial workshop</td>
<td></td>
</tr>
<tr>
<td>• Initial Drafting</td>
<td>November 2017</td>
<td>February 2019</td>
</tr>
<tr>
<td>December 2017</td>
<td>• Facilitator update workshop</td>
<td>May 2018</td>
</tr>
<tr>
<td>• Finalize program</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There is a wide range in type and severity of symptoms among children affected by Zika. It was agreed that program inclusion criteria would be:

- Caregivers of children who have confirmed or suspected CZS but not other types of neurodevelopmental disabilities
- Caregivers of children residing at home and not currently requiring inpatient hospital care
Caregivers willing to attend the whole program and living within 1 hour of the group meeting location

Children of any age (although given the nature of the epidemic in Brazil in 2017–2018, they were all aged 3 years and younger)

Children who may be receiving rehabilitation services to address individual needs

More than 1 caregiver of a child (e.g., mother and father, mother and grandmother) was allowed to attend the group meeting.

From June to August 2017, the project lead researcher (AD) drafted the initial program, adapting the GTKCP and EIP materials with input from expert committee members, project teams in Brazil, and other experts (TS, MS, SF, EM, HK).

**PILOTING**
We piloted the approach during 2 phases with 6 different groups and used this information to finalize the program. A future analysis will report the feasibility of the program using qualitative and quantitative data analysis.

**Program Establishment in Brazil**
The partners in Brazil (IFF and UFBA) identified a site coordinator (MS and SF) for each of the 2 pilot sites, the states of Rio de Janeiro and Bahia. The site coordinators’ main responsibilities were to manage the logistic components of the pilot groups, including identifying an appropriate location for the groups, recruiting facilitators, recruiting researchers, identifying participants, and liaising with local health providers.

Rio de Janeiro and Greater Salvador, Bahia, were selected as pilot sites because they had a large population of children impacted by Zika. Recife, which the LSHTM team visited during the initial country visit, was not selected because several other intervention projects by other organizations were already taking place and contamination of outcomes was a concern. Three sites within Rio de Janeiro and 3 municipalities of Greater Salvador (Simões Filho, Lauro de Freitas, and Camaçari) were selected because of their proximity to families of children with CZS, availability of an appropriately sized venue, and willingness of the local relevant authorities to accommodate a group.

Facilitators were identified by the site coordinators and approved by the project team lead. A total of 8 local facilitators were selected (4 therapists with experience in pediatrics and CZS and 4 expert mothers). In August 2017, a week-long facilitator training was conducted in Rio de Janeiro and led by a trainer who has taught the GTKCP program extensively. The trainer was
Program Development for Families Impacted by Congenital Zika Syndrome

Two additional support groups were established in each pilot setting (4 total), with the primary aim of ascertaining the feasibility of the intervention. These support groups had identical procedures for data collection, real-time feedback, and fast-track learning (February–June 2018). After the delivery of the groups, the intervention was further updated, improved, and finalized using the same processes as before. The 2 groups in Rio had 7 and 9 families, respectively, and the 2 groups in Greater Salvador had 10 and 7 families, respectively.

Summary
Six groups ran between August 2017 and June 2018 across 2 phases. The children of the caregivers were 25 males and 23 females with an average age of 23 months (standard deviation=9 months) at their first session. Of the families included in all 6 pilot groups, all (n=48) stated the mother as the primary caregiver. The ages of the mothers (n=48) were 15–20 years (3), 21–25 years (17), 26–30 years (5), 31–40 years (18), and 41–50 years (3). Thirty-six mothers reported they were married, 3 divorced, and 9 reported they were single. Only 6 mothers reported being in work, with the most common reason for not being in work being that they cared for their child (n=34).

During the second and third groups in Greater Salvador, held between January and June 2018, several children with non-Zika related developmental disabilities participated in the sessions. This was done for 2 reasons: (1) to increase the number of children participating because the number of children with CZS who met the inclusion criteria was quite low, and (2) to assess whether combining caregivers of children with CZS and those with other neurodevelopmental disabilities would be a positive experience.

We focused primarily on the caregiver and the program, with some interaction with the family, community, and services at the activity and output levels as informed by our theory of change (Figure 2). The proximal outcomes of the program are expected to be (1) increased participant quality of life and confidence in caring for a child with CZS, and (2) an intervention that is feasible to scale up and replicate in other contexts. Core to the theory of change is empowering the caregiver to improve care for their child through developing support networks and increased knowledge and awareness of their child’s needs.

Pilot Phase 1
In August 2017, the first 2 pilot support groups—1 in Rio de Janeiro and 1 in Greater Salvador—started meeting weekly. The Rio group had 7 families, and the Greater Salvador group had 8. There were 10 sessions for each group with a different topic each week. Researchers used 3 techniques to collect data to inform real-time feedback and fast-track learning about the content and processes of the session. First, researchers directly observed the sessions and noted the session flow, participants’ responses, and behaviors of participants and facilitators. Second, researchers conducted focus group discussions at the end of each of the 10 sessions with participants and (separately) with facilitators to obtain immediate reflections and feedback on the session content. The researchers recorded detailed observation notes about the session and comprehensive notes about focus group discussions that they uploaded to a password-secured Google Drive document for the content developer (TS) to analyze. Third, researchers recorded pertinent comments from participants, facilitators, and site coordinators on images, content, activities, practicalities, and logistics, which were made outside of the sessions. Weekly calls within 48 hours of the session occurred between the researchers and TS, which allowed for further explanation and contextualization. Content issues were recorded and reviewed to update the program in real time and for 4 weeks after the conclusion of phase 1 in November 2017.

Pilot Phase 2
In December 2017, a 3-day training session provided facilitators and site coordinators with information on the changes to the program content and structure based on fast-track learning in the first pilot phase.

Two researchers were identified by the site coordinators and approved by the project team lead. All the researchers had a background in psychology, but this was not a prerequisite for the role. The researchers participated in a 2-day training in July 2017 on the research approaches and data collection methods and on the fast-track learning approach that would be used to update and adjust the program content based on weekly feedback that they collected from the groups.

In November 2017, the Rio group started meeting weekly. The Rio group had 7 families and the Greater Salvador group had 8. There were 10 sessions for each group with a different topic each week. Researchers used 3 techniques to collect data to inform real-time feedback and fast-track learning about the content and processes of the session. First, researchers directly observed the sessions and noted the session flow, participants’ responses, and behaviors of participants and facilitators. Second, researchers conducted focus group discussions at the end of each of the 10 sessions with participants and (separately) with facilitators to obtain immediate reflections and feedback on the session content. The researchers recorded detailed observation notes about the session and comprehensive notes about focus group discussions that they uploaded to a password-secured Google Drive document for the content developer (TS) to analyze. Third, researchers recorded pertinent comments from participants, facilitators, and site coordinators on images, content, activities, practicalities, and logistics, which were made outside of the sessions. Weekly calls within 48 hours of the session occurred between the researchers and TS, which allowed for further explanation and contextualization. Content issues were recorded and reviewed to update the program in real time and for 4 weeks after the conclusion of phase 1 in November 2017.

Pilot Phase 1
In August 2017, the first 2 pilot support groups—1 in Rio de Janeiro and 1 in Greater Salvador—started meeting weekly. The Rio group had 7 families, and the Greater Salvador group had 8. There were 10 sessions for each group with a different topic each week. Researchers used 3 techniques to collect data to inform real-time feedback and fast-track learning about the content and processes of the session. First, researchers directly observed the sessions and noted the session flow, participants’ responses, and behaviors of participants and facilitators. Second, researchers conducted focus group discussions at the end of each of the 10 sessions with participants and (separately) with facilitators to obtain immediate reflections and feedback on the session content. The researchers recorded detailed observation notes about the session and comprehensive notes about focus group discussions that they uploaded to a password-secured Google Drive document for the content developer (TS) to analyze. Third, researchers recorded pertinent comments from participants, facilitators, and site coordinators on images, content, activities, practicalities, and logistics, which were made outside of the sessions. Weekly calls within 48 hours of the session occurred between the researchers and TS, which allowed for further explanation and contextualization. Content issues were recorded and reviewed to update the program in real time and for 4 weeks after the conclusion of phase 1 in November 2017.

Pilot Phase 2
In December 2017, a 3-day training session provided facilitators and site coordinators with information international, and we used a translator for the sessions as well as materials in Brazilian Portuguese. The training involved education on facilitating a group, practice sessions with reflective learning and feedback, and opportunities for discussion. The project leads and site coordinators selected 2 pairs of facilitators to lead the first pilot groups based on their performance during the training week.

Two researchers were identified by the site coordinators and approved by the project team lead. All the researchers had a background in psychology, but this was not a prerequisite for the role. The researchers participated in a 2-day training in July 2017 on the research approaches and data collection methods and on the fast-track learning approach that would be used to update and adjust the program content based on weekly feedback that they collected from the groups.

Pilot Phase 1
In August 2017, the first 2 pilot support groups—1 in Rio de Janeiro and 1 in Greater Salvador—started meeting weekly. The Rio group had 7 families, and the Greater Salvador group had 8. There were 10 sessions for each group with a different topic each week. Researchers used 3 techniques to collect data to inform real-time feedback and fast-track learning about the content and processes of the session. First, researchers directly observed the sessions and noted the session flow, participants’ responses, and behaviors of participants and facilitators. Second, researchers conducted focus group discussions at the end of each of the 10 sessions with participants and (separately) with facilitators to obtain immediate reflections and feedback on the session content. The researchers recorded detailed observation notes about the session and comprehensive notes about focus group discussions that they uploaded to a password-secured Google Drive document for the content developer (TS) to analyze. Third, researchers recorded pertinent comments from participants, facilitators, and site coordinators on images, content, activities, practicalities, and logistics, which were made outside of the sessions. Weekly calls within 48 hours of the session occurred between the researchers and TS, which allowed for further explanation and contextualization. Content issues were recorded and reviewed to update the program in real time and for 4 weeks after the conclusion of phase 1 in November 2017.

Pilot Phase 2
In December 2017, a 3-day training session provided facilitators and site coordinators with information
Fast-track learning meant that the intervention was updated and improved as new information was gathered each week about what was working or not. For example, practical or administrative issues, such as organization of transport for participants, were changed and updated in real time each week.

As a result of rapid participant feedback, we made several changes to the program. For example, we changed the title of session 8 (highlighting advocacy and empowerment) to “uniting our voices”; the original title “raising our voices” translated to “shouting out loud.” In a second example, participants felt that the images used in the first 2 pilots, which used images from GTKCP and EIP, did not adequately reflect phenotype, family behavior, and environment in Brazil. Therefore, as participants requested, we included images that reflected their lives to create identification and favor more adherence. A local artist was engaged to draw more culturally appropriate images for the later groups, which were perceived more positively. More representation of fathers in caring roles was also incorporated at this stage.

New innovations in Juntos, which were not in GTKCP or EIP, include information on the Zika virus, strengthened participatory approaches to engage participants with community inclusion and disability rights, and a concerted effort to improve male engagement, which was successful to a degree (though the female engagement was still much higher). Additional content includes group discussion on gastrostomy (dysphagia was a common problem), creating trousers stuffed with padding to support children in sitting, using an elasticated cloth to rock children who are irritable, and activities to promote understanding of disability rights. In addition, each session includes reflection and discussion on the session and on the past week through an emotional support activity at the end of the session. The facilitators work as a pair throughout the session; however, the emotional support activity is facilitated by the expert mother. The first 5 sessions include the same activity with facilitated questions:

- How did you find talking about today’s subject?
- Did it raise any emotions or feelings that you did not expect?
- How have you been feeling this week?

The predictability of the questions helps participants to become comfortable with sharing. By week 5, participants have explored much of their thoughts on emotions and feelings, and this then progresses to reflecting on the future.

The feasibility assessment is not detailed in this article and will be described in a future article on the findings.

Finalization of the Program
Consensus on the final content of the program was reached through 2 workshops (London, United Kingdom, and Rio de Janeiro, Brazil) in May 2018. One group in Greater Salvador was still running. However, feedback that had already been collected from the groups was deemed sufficient to be able to finalize the content. The workshops included the technical advisory committees, study site coordinators, and researchers (psychologists).

## PROGRAM DESCRIPTION

The final program intervention is called Juntos, which means together in Portuguese and Spanish, to emphasize the importance of inclusion and mutual support. Intervention materials comprise a facilitator manual and participant materials, such as photographs, animations, and video footage. An allied health professional and an expert mother cofacilitate groups that meet once a week for 10 sessions. Support and guidance for facilitators is provided by project coordinators via telephone, email, and/or WhatsApp.

Groups are held at local community facilities, such as health centers, offices of local organizations, or schools, to minimize participants’ travel time and to foster relationships between people who lived relatively near to each other. Nine sessions are only for the caregivers and their children, and 1 session is open for other community members to attend. The children who come are looked after in a separate room or space by volunteers, but they are present for some of the practical aspects whenever relevant. Table 1 describes each Juntos module.

The sessions are participatory and use principles of adult learning theory. Participants learn by sharing their own experiences and realities about topics that are important to them, which promotes peer support, critical thinking, and mutual problem solving. The groups start with a light-hearted icebreaker to welcome and warm up the conversation and to encourage comfortable interaction. Participants are then guided through a series of activities, open discussions, pair work, explanations, and demonstrations. Tables 2 and 3 provide examples of session content from session 4 and 6, respectively.

Supportive information was developed for the program that includes short videos on the program
and different aspects of care. The individual modules, full manual, and supportive materials are available in English, Portuguese, and Spanish: https://www.ubuntu-hub.org/resources/juntos.

LESSONS LEARNED

Fast-track learning added value to the intervention development because it allowed inclusion of language, logistics, content, and culturally specific changes in real time. Participants’ feedback during the first pilot phase was utilized to revise the content (for example, providing case studies, images, and videos of fathers undertaking practical tasks), which may have made the overall content more useful for the later groups. The later groups were aware of this process and recognized some of the changes based on early peers’ feedback. In a context of relative distrust and research fatigue, this process helped to demonstrate how participant feedback was valued and reinforced that the program was genuinely and specifically intended for caregivers, a group that has been overlooked in the wider Zika response.

The integration of a component of caregiver emotional well-being in this group intervention demonstrates a novel approach to including psychosocial support to better promote emotional well-being as an integral part of health work, rather than being seen as a standalone effort. There is no single recognized theory of how participatory groups achieve their health impacts and few studies evaluate how and why different support networks improve caregiver and child outcomes. Examples in resource-limited settings include self-help groups for people with mental health conditions, which demonstrate positive impacts on both the people with mental health conditions and their caregivers. Additionally, women’s self-help groups have resulted in improved maternal and neonatal survival. Our integration of a mental health component in Juntos illustrates that groups that address child development can practically integrate emotional support activities. Facilitators reported that they valued having a dedicated space each week to raise issues of emotional well-being. The practical components of the sessions often raised some emotions for a participant, but there would be little time to explore these, so the final section allowed further exploration and discussion between the group. Evaluation of whether such a strategy can work in other settings is necessary, and negative and unanticipated consequences warrant further evaluation in future work. Having an expert mother facilitate these sessions was particularly important and helped form group connections that might not have been possible with an allied health professional alone.

In understanding pathways to change, the role of the expert mother appears to offer crucial...
encouragement to shared learning between caregivers and contributes to developing an egalitarian atmosphere, expanding care practices beyond traditional rehabilitation models. Relating this common ground and a sense of belonging through a social support network provides an environment to improve the knowledge and skills of caregivers. It was critically important that the 2 cofacilitators were equals, each bringing their own experiences to the process and an expertise and insight that the other did not possess. The allied health professionals immediately saw the value in this, and there was no sense of protectionism or defensiveness that they needed to be the lead or expert given their professional training.

Groups were held in the local community so that caregivers could build strong local networks. This also increased interest from caregivers of

---

**TABLE 1. Finalized Module Topics Included in Juntos, A Community-Support Group for Caregivers of Children with Congenital Zika Syndrome in Brazil**

<table>
<thead>
<tr>
<th>Module</th>
<th>Topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Introduction</td>
<td>• About the program</td>
</tr>
<tr>
<td></td>
<td>• Information about Zika and Congenital Zika Syndrome</td>
</tr>
<tr>
<td></td>
<td>• How to find information</td>
</tr>
<tr>
<td></td>
<td>• Personal stories</td>
</tr>
<tr>
<td>2. Our child</td>
<td>• Introducing your close family and friends</td>
</tr>
<tr>
<td></td>
<td>• Development milestones for young children</td>
</tr>
<tr>
<td></td>
<td>• Determining your child’s progress</td>
</tr>
<tr>
<td></td>
<td>• Managing irritability and crying</td>
</tr>
<tr>
<td>3. Positioning and moving</td>
<td>• How to position children who need assistance</td>
</tr>
<tr>
<td></td>
<td>• How to assist children to learn to move</td>
</tr>
<tr>
<td>4. Eating and drinking</td>
<td>• Feeding challenges</td>
</tr>
<tr>
<td></td>
<td>• Practical skills to address challenges for your child</td>
</tr>
<tr>
<td>5. Communication</td>
<td>• Importance of communication</td>
</tr>
<tr>
<td></td>
<td>• Practical advice to help your child communicate</td>
</tr>
<tr>
<td>6. Play and early stimulation</td>
<td>• Importance of play for children to develop and learn</td>
</tr>
<tr>
<td></td>
<td>• Early stimulation</td>
</tr>
<tr>
<td></td>
<td>• Making simple toys</td>
</tr>
<tr>
<td></td>
<td>• Inclusion of play in the family and broader community</td>
</tr>
<tr>
<td>7. Everyday activities</td>
<td>• How to use everyday activities to help your child develop</td>
</tr>
<tr>
<td></td>
<td>• Managing seizures</td>
</tr>
<tr>
<td>8. Uniting our voices</td>
<td>• Understand the context of disability rights</td>
</tr>
<tr>
<td></td>
<td>• Education</td>
</tr>
<tr>
<td></td>
<td>• Communicating with your health team</td>
</tr>
<tr>
<td></td>
<td>• Advocating</td>
</tr>
<tr>
<td>9. Our community</td>
<td>• Who is in your community?</td>
</tr>
<tr>
<td></td>
<td>• Common barriers to inclusion</td>
</tr>
<tr>
<td></td>
<td>• Addressing negative attitudes and exclusion</td>
</tr>
<tr>
<td></td>
<td>• Social activity</td>
</tr>
<tr>
<td>10. Next steps</td>
<td>• Summing up</td>
</tr>
<tr>
<td></td>
<td>• Planning next steps for yourself and the group</td>
</tr>
</tbody>
</table>
children with developmental disabilities other than CZS and highlights the importance of de-isolating Zika from other causes of neurodevelopmental disability when developing community support programs. Juntos does not replace health care services but rather seeks to complement services by empowering other caregivers to optimize their child’s care and upbringing.

We received positive feedback during the sessions that combined caregivers of children with CZS with caregivers of children with other neurodevelopmental disabilities. There was a recurrent expression of comfort among the caregivers when engaging with other caregivers in similar situations and circumstances that they were not as alone, unique, and isolated as they had perhaps feared. This was also seen in the sessions where non-CZS caregivers engaged and, in fact, there was a value perceived to understand that the challenges being faced were not unique to only caregivers of CZS. This was also reinforced frequently in session 8 of the Rio sessions, where an external speaker came from a local Down’s Syndrome organization to discuss their advocacy approaches; the sessions were always extremely well received by participants. Although the challenges facing children with CZS and their caregivers remain unique and, to a certain degree, still unknown, there may be an important value to ensure that there are also many common issues faced and a shared approach may be both efficient and useful.

---

**TABLE 2. Example of Content From Facilitated Group Session 4 on Eating and Drinking from Juntos, A Community-Support Group for Caregivers of Children With Congenital Zika Syndrome in Brazil**

<table>
<thead>
<tr>
<th>Example</th>
<th>Discussion</th>
<th>Aim</th>
</tr>
</thead>
<tbody>
<tr>
<td>Icebreaker</td>
<td>In pairs: One person tries to give the other a drink of water in different positions (e.g., head leaning back, turned to one side, or flopping forwards)</td>
<td>How easy or difficult is it to swallow in each position? How does it feel to be fed?</td>
</tr>
<tr>
<td>Discussion</td>
<td>As a large group to share experiences</td>
<td>What is a nutritious or “balanced” diet?</td>
</tr>
<tr>
<td>Activity</td>
<td>Show a banana and a biscuit and other common food</td>
<td>Discuss—Are the items hard or soft? Can they be made into a smooth puree? How?</td>
</tr>
</tbody>
</table>

**TABLE 3. Example of Content From Facilitated Group Session 6 on Play and Early Stimulation from Juntos, A Community-Support Group for Caregivers of Children With Congenital Zika Syndrome in Brazil**

<table>
<thead>
<tr>
<th>Example</th>
<th>Discussion</th>
<th>Aim</th>
</tr>
</thead>
<tbody>
<tr>
<td>Icebreaker</td>
<td>In groups of 3: each group is given one inexpensive everyday item (e.g., cup, piece of cloth, container, ball) and everyone uses their imagination to transform the object into something else and acts it out</td>
<td>What is play?</td>
</tr>
<tr>
<td>Discussion</td>
<td>As a large group to share experiences</td>
<td>What have you found play helps your child to do? Does your child need to play?</td>
</tr>
<tr>
<td>Activity</td>
<td>Toy making, such as making bells and rings with ribbons</td>
<td>Discuss—How can you involve short periods of play in your daily activities? How can you involve other members of your family in playing with your child?</td>
</tr>
</tbody>
</table>
RECOMMENDATIONS

The needs analysis that we undertook at the beginning of the project\(^2\) as well as more recent literature\(^27,28\) has highlighted overlaps and similarities between cerebral palsy and CZS. We suggest that children with CZS and their caregivers may benefit by integrating and linking with services and programs for children with other neurodevelopmental disabilities. Rehabilitation/therapy services were already doing this to a large extent, and there seems a good scope for other health and social service providers to also ensure service integration. Conversely, newly formed services as a result of the attention to CZS shouldn’t be exclusive to this population group and should seek to include all families and children who may benefit.

By the nature of its design, Juntos can potentially be implemented by a range of stakeholders, from nongovernmental organizations to public community services to primary health settings. This flexibility may mean that there is a stronger opportunity for Juntos to be scaled up. The universal primary health structure in Brazil—the Sistema Único de Saúde—could be an avenue to further explore. We see opportunities for public/private partnerships also. Cost is clearly a major factor in the potential for scale up. Facilitator training can be done in larger groups to reduce costs. In addition, if the facilitator therapists undertake the role as part of their existing work, these costs may be further reduced. However, we do feel that it is important to remunerate parent facilitators for their work and other costs, such as transport and refreshments, to ensure full participation of families.

Strengths

Strengths of this pilot include the development process being informed by a theory of change and reflective practice and robust methodology that allowed integration of rapid feedback. Real-time feedback and adaption enabled the development of a culture-specific and language-specific intervention, and the program was developed and refined to meet the needs of caregivers of children with CZS in Brazil. Running the program in 2 sites concurrently (Rio de Janeiro and Greater Salvador) was an important methodological choice for achieving better final version program. Brazil is huge and diverse, and although these 2 sites do not cover the breadth of diversity, piloting in more than 1 site and acquiring different feedback added to the strength of the study.

Limitations

Our study has limitations. We describe the intervention development, but assessment of feasibility and evaluation of replication and scale-up in other countries is now needed. More work is needed on forming a comprehensive facilitator training program, and further development of the intervention to include all children with developmental disabilities is warranted. If Juntos is found to be feasible, robust studies to evaluate the cost-effectiveness of the intervention will be needed.

CONCLUSIONS

We developed and refined a participatory community-based group intervention to meet the needs of caregivers of children with CZS. Juntos has the potential to be an important resource for community practice. There is scope to expand across Brazil and in other South American countries and to children with other developmental disabilities.

Acknowledgments: The authors wish to acknowledge the valuable insights and inputs made by many people in the UK, Brazil and elsewhere in supporting the development of the program including Mel Adams, Julian Eaton, Maria Antônia Goulart, Mila Mendonça, Kate Milner, Cally Tann, Liana Ventura, Joerg Weber, Maria Zuurmond.

Funding: The project was funded by Wellcome Trust and Department for International Development, Grant Number 206719/12/17/2.

Competing interests: AD joined the Pan American Health Organization (PAHO) during the research period. Work on the research study was undertaken outside and separate to his PAHO duties.

REFERENCES

Juntos: Um Programa de Apoio às Famílias Afetadas pela Síndrome Congênita do Vírus Zika no Brasil

ABSTRACTO

Histórico: O surto, sem precedentes, do vírus do Zika em 2015–2016 resultou no nascimento de mais de 3.000 crianças com a Síndrome Congênita do Vírus Zika (SCZ). Essas crianças experenciam múltiplas e complexas condições de saúde com limitado acesso a serviços de apoio tanto para elas quanto para as suas famílias.

O desenvolvimento de um programa piloto: um programa existente de apoio às crianças com paralisia cerebral (Getting to Know Cerebral Palsy) foi adaptado para o contexto do Zika no Brasil através de uma consultoria especializada. O programa piloto foi testado em dois locais com 6 grupos de cuidadores (um total de 48 famílias) entre agosto de 2017 e junho de 2018. Grupos focais e de observação – Conhecendo a Paralisia Cerebral) foi adaptado para o contexto do Zika no Brasil através de uma consultoria especializada. O programa piloto foi testado em dois locais com 6 grupos de cuidadores (um total de 48 famílias) entre agosto de 2017 e junho de 2018.

Descrição do programa: a iniciativa Junto é um programa de facilitação e participação para grupos de cuidadores de crianças com SCZ. São dez encontros com duração de 4 horas - cada um inclui uma dinâmica inicial de quebra-gelo, atividades e discussões em grupo. O conteúdo cobre informações práticas sobre os cuidados com crianças com problemas de desenvolvimento, incluindo aqueles causados pelo Zika. O apoio psicossocial abrange um componente importante no qual as famílias são orientadas desde a primeira semana sobre como definir e desenvolver suporte em suas comunidades. A realização com seis grupos no Rio de Janeiro e na aérea metropolitana de Salvador ocorreu de forma exitosa e em ambos os locais houve um retorno positivo em termos de aceitação e demanda.

Conclusões: o programa tem o potencial de ser uma ferramenta importante para as aéreas de saúde e prestação de serviços sociais na América do Sul em resposta ao vírus do Zika. Além disso, pode ser adaptado para crianças com problemas no neurodesenvolvimento para além daqueles causados pelo Zika, o que por sua vez, pode ser importante para garantir que as famílias de crianças com SCZ sintam-se menos isoladas.
Aspectos Principais

- Um programa de orientação comunitária para grupos de família desenvolvido para os cuidadores de crianças com a Síndrome Congênita do vírus Zika (SCZ), baseado em programa anterior focado em crianças com paralisia cerebral, foi testado como uma experiência piloto no Brasil.
- Gestores que desenvolvem atividades de grupo devem considerar as abordagens de aprendizagem rápida para adequar as suas intervenções, tornando-as mais úteis para os participantes.
- Médicos e gestores de iniciativas para o apoio às vítimas do Zika devem atentar para o fato de que cuidadores de crianças com CZS podem se beneficiar da interlocução e troca com cuidadores de crianças com outras problemas de neurodesenvolvimento.
- Gestores de programas focados em intervenções comunitárias devem levar em consideração engajar as mães como um grupo de cofacilitadoras.
- Integrar atividades de apoio emocional em grupos para abordar a questão do desenvolvimento infantil é algo prático importante e agrega valor.

Peer Reviewed

Received: December 23, 2019; Accepted: November 4, 2020


© Duttine 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-20-00018
Improve Hospital Oxygen Systems for COVID-19 in Low-Resource Settings: Lessons From the Field

Hamish R. Graham, Sheillah M. Bagayana, Ayobami A. Bakare, Bernard O. Olayo, Stefan S. Peterson, Trevor Duke, Adegoke G. Falade

Key Messages
- The COVID-19 pandemic has highlighted and exacerbated deficiencies in hospital oxygen systems globally but is also an opportunity to “build back better.”
- Our collated field experience from African and Asia-Pacific contexts reveal practical strategies whereby hospitals can rapidly improve their oxygen systems. We share guidance documents (all open access) for local use and adaptation.

Key Implications
- Using our practical guides, hospital staff can help:
  - Improve pulse oximetry and oxygen use
  - Optimize existing oxygen supplies
  - Expand existing oxygen systems with robust equipment and smart design
- Policy makers and program managers can use our recommendations to ensure that investments in oxygen systems are deployed and implemented more effectively and efficiently.

ABSTRACT
Oxygen therapy is an essential medicine and core component of effective hospital systems. However, many hospitals in low- and middle-income countries lack reliable oxygen access—a deficiency highlighted and exacerbated by the coronavirus disease (COVID-19) pandemic. Oxygen access can be challenged by equipment that is low quality and poorly maintained, lack of clinical and technical training and protocols, and deficiencies in local infrastructure and policy environment. We share learnings from 2 decades of oxygen systems work with hospitals in Africa and the Asia-Pacific regions, highlighting practical actions that hospitals can take to immediately expand oxygen access. These include strategies to: (1) improve pulse oximetry and oxygen use, (2) support biomedical engineers to optimize existing oxygen supplies, and (3) expand on existing oxygen systems with robust equipment and smart design. We make all our resources freely available for use and local adaptation.

BACKGROUND
Oxygen therapy is an essential medicine and core component of hospital systems that has been a standard of care for more than 100 years. However, access to oxygen therapy is limited in many low-resource settings, where the majority of hypoxemic patients who are admitted to the hospital will not receive oxygen, resulting in an increased risk of death.

The coronavirus diseases (COVID-19) pandemic has revealed the extent of this “oxygen gap” and stimulated long overdue interest in improving oxygen systems. Approximately 20% of patients who have COVID-19 require hospital admission for oxygen therapy (with or without extra respiratory support). Although much attention has focused on ventilator and intensive care unit capacity, improving basic hospital oxygen systems must take priority.

CHALLENGES IN OXYGEN ACCESS
To provide oxygen therapy, we need a reliable oxygen supply, prompt identification of patients requiring oxygen therapy, and appropriate administration by skilled health care workers. Oxygen supply is typically achieved using oxygen cylinders (filled at an oxygen plant), oxygen concentrators (concentrating oxygen from air on-site), oxygen plants (piped directly or distributed via cylinders), or liquid
oxygen (delivered from a specialized gas plant and stored on-site at very high pressure). Oxygen use is guided by nurses measuring blood oxygen levels using a pulse oximeter (or relying on clinical signs if no oximeter is available) and titrating oxygen flow rate to maintain adequate blood oxygen levels.

However, achieving reliable supply and appropriate use is challenging, with major barriers due to equipment that is low quality and poorly maintained, lack of clinical and technical training and protocols, and deficiencies in local infrastructure and sociopolitical context.6

For example, surveys in Nigeria have found that although half of hospitals had oxygen cylinders or concentrators on inpatient wards, the cylinders and concentrators were frequently empty or nonfunctional.2,7 Detailed testing in a selection of these hospitals found that only 5% of concentrators tested were producing medical grade oxygen.2 Almost no hospitals had pulse oximeters available on the wards.2,7 Procurement of oxygen equipment was haphazard, preventive maintenance was nonexistent, and hospital technicians were untrained and under-supported. Hospital nurses were unfamiliar with pulse oximetry, and the majority of hypoxic patients were not receiving oxygen.2,7 Hospital directors bemoaned the cost of oxygen, with one director describing oxygen as his “biggest headache.”

We, and others, have reported similar findings in Kenya, Uganda, Papua New Guinea, and other African and Asia-Pacific contexts.7–11 Indeed, unreliable oxygen supplies and deficiencies in oxygen use are consistent and persisting problems for many hospitals in these regions, particularly in rural and remote settings.

Challenges to oxygen access exist alongside broader systems issues such as unreliable power supply, health care workforce constraints, high out-of-pocket health care costs, low health literacy, and underfunded public health and preventive services.

However, our work has shown that effective and sustainable change is possible.

LESSONS LEARNED

Over the past 2 decades, we have supported hospitals in Africa and Asia-Pacific regions to improve oxygen systems using low-cost technology such as oxygen concentrators, pulse oximeters, and cylinder distribution systems. Our work has shown how to combine quality equipment and training to achieve context-appropriate and sustainable improvement in oxygen systems and improve clinical outcomes.5,11–13 This experience has informed the development of clinical and technical guidance by the World Health Organization (WHO) and the United Nations Children’s Fund (UNICEF),14–16 and is now informing oxygen scale-up for COVID-19 (Box).

To complement existing technical guidance on COVID-19 response from WHO17 and others, we offer practical suggestions based on our on-the-ground experience to help policy makers, administrators, technicians, and health care workers seeking to rapidly improve their hospital oxygen systems.

1. Support Health Care Workers to Use Pulse Oximetry and Oxygen Through Training and Protocols

In many low- and middle-income countries (LMICs), oxygen is absent from medical and

---

**BOX. Essential Resources for Additional Information on Oxygen and COVID-19**

- WHO Academy’s COVID-19 mobile learning app for Android and iPhone/iPad, which contains much of the clinical and technical advice in a conveniently accessible format.
nursing training,¹⁸ and pulse oximetry is unavailable.² Pulse oximetry, with practical task-based training and simple guidelines, can enable health care workers to target oxygen toward those who need it most, dramatically improving patient oxygen access and clinical outcomes.¹³ In contexts where pulse oximetry is not a standard of care, COVID-19 offers an opportunity to establish oxygen saturation as the “fifth vital sign.” However, although pulse oximetry is a simple skill, it is a fundamentally new concept for many health care workers and will require encouragement and support to integrate it into the workflow.¹⁹ Education and support for health care workers to use pulse oximetry and oxygen therapy should also cover maintenance and functioning of oxygen equipment, recognizing the critical importance of health care worker and technician teamwork in maintaining a reliable oxygen supply. We have created practical training materials,
2. Assist Biomedical Engineers to Optimize Existing Oxygen Supplies Through Training, Protocols, and Logistic Support

Oxygen is a medicine that depends on technology and requires effective teamwork between health care workers, technicians, and managers. However, biomedical engineers and hospital technicians are frequently left out of decision-making processes and lack maintenance budgets or system support. Training, provision of tools and spare parts, and stronger maintenance and transport systems can enable repair and optimization of existing oxygen equipment and supply chains. Installation of simple piping and individual flowmeters can improve safety (allowing individual titration of flow), efficiency (sharing a single oxygen source between multiple patients) and infection control (allowing oxygen sources to be kept away from patient areas). Including technicians alongside health care workers in multidisciplinary teams can help transform a problem-driven “oxygen headache” into focused oxygen solutions. We have created practical resources to assist biomedical engineers/technicians to build and maintain reliable, user-friendly oxygen systems using oxygen concentrators and/or cylinders, flowmeter stands, and simple piping (Supplement 4).

3. Expand on Existing Oxygen Systems Using Robust Equipment and Smart Design

WHO and UNICEF have released guidance on oxygen-related equipment and specific guidance for COVID-19. This guidance includes low-cost oxygen concentrator-based systems that use simple plastic piping and flowmeter stands to provide oxygen to multiple patients simultaneously. These systems have been successfully implemented in African and Asia-Pacific contexts and can be established in a relatively short time frame (compared to a new oxygen plant). With the support of several other donors, UNICEF has delivered almost 15,000 oxygen concentrators and approximately 15,000 pulse oximeters to 69 countries (at the time of writing). Many other donors have channeled equipment support directly. However, there is a real risk that these valuable investments will end up in equipment graveyards with inadequate consideration to how they are deployed in hospitals. Hospitals can use our practical installation guidance to create smart and efficient ward oxygen systems to put this influx of equipment to use rapidly and effectively (Supplement 4).
CONCLUSIONS

Improving patient outcomes always hinges on doing the basics well. The COVID-19 pandemic offers the opportunity to refocus efforts on the basics of acute care, knowing that improvements in oxygen (as well as infection control, triage, laboratory testing, etc.) will benefit patients both now and in the future.

Improving oxygen systems is an achievable priority for hospitals in LMICs. We propose practical steps to support effective and sustainable improvements in hospital oxygen systems during the COVID-19 pandemic. We share these learnings in the hope that health care workers, techni-
cians, hospital managers, and policy makers will be able to take immediate actions toward better oxygen access. All the accompanying oxygen resources are freely available for users to download, use, and adapt to local needs (https://bit.ly/OxygenResources). We welcome your feedback.

Acknowledgments: We would like to thanks Bev Bradley, Cindy McWhorter, and the United Nations Children’s Fund supply division technical team; Janet Diaz and the team at the World Health Organization; and Olatayo Olatinwo and the Oxygen for Life Initiative team.

Competing interests: HG, AAB, TD, AGF received funding from the Bill and Melinda Gates Foundation outside the submitted work. The funders had no role in the writing or submission of this article. All other authors declare no competing interests.

REFERENCES


Peer Reviewed

Received: May 18, 2020; Accepted: September 21, 2020; First published online: September 28, 2020


© Graham 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-20-00224
Corrigendum: Parmaksiz K et al., What Makes a National Pharmaceutical Track and Trace Succeed? Lessons From Turkey

In the article “What Makes a National Pharmaceutical Track and Trace Succeed? Lessons From Turkey” by Koray Pamaksiz et al. (Volume 8, Issue 3), the Funding statement on page 10, incorrectly listed the U.S. Pharmacopeia Quality Institute as one of the funders of the work. The article has been corrected accordingly.


© Global Health: Science and Practice. This is an open-access article distributed under the terms of the Creative Commons Attribution License, 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-20-00587