EDITORIALS

Human-Centered Design and Sustainable Malaria Interventions

Human-centered design provides a method to adapt malaria control interventions to be more closely aligned with a family’s convenience, comfort, and personal lifestyle, enabling a broader and more sustained culture of access and use.

Michael Macdonald, Thomas Putzer
https://doi.org/10.9745/GHSP-D-19-00189

COMMENTARIES

Supervision of Task-Shared Mental Health Care in Low-Resource Settings: A Commentary on Programmatic Experience

Task-shared mental health care programs in low-resource settings often incorporate supervisory structures that would be difficult to implement at scale, and many rely on foreign specialist experts as supervisors. Future programs could leverage peer supervision, technology, competency assessments/fidelity checklists, and other tools. Mental health care specialists will require training, support, and incentives to supervise generalist care providers.

Christopher G. Kemp, Inge Petersen, Arvin Bhana, Deepa Rao
https://doi.org/10.9745/GHSP-D-18-00337

ORIGINAL ARTICLES

Using a Human-Centered Design Approach to Determine Consumer Preferences for Long-Lasting Insecticidal Nets in Ghana

Through focus group discussions and human-centered design exercises, middle-class Ghanaians communicated the need to address convenience, comfort, and aesthetics when designing a bed net for their demographic. Illustrative attributes for consideration by private-sector manufacturers include a more convenient way to hang the net, a more attractive silhouette, and a zipper to provide ease of entry and exit while keeping the area sealed from mosquitos.

Sharon Kim, Danielle Piccinini, Elorm Mensah, Matthew Lynch
https://doi.org/10.9745/GHSP-D-18-00284
Cell Phone Counseling Improves Retention of Mothers With HIV Infection in Care and Infant HIV Testing in Kisumu, Kenya: A Randomized Controlled Study

Tailored, one-on-one counseling delivered via cell phone was very effective in retaining mothers with HIV in care and in promoting infant HIV testing and antenatal and postnatal care attendance. The highest risk of loss to follow-up among women with HIV accessing PMTCT services was prior to delivery and then after infant HIV testing at 6 weeks. Challenges include continued limited access to cell phones, difficulty with reaching participants on the phone, and poor adherence to antiretroviral therapy for a substantial percentage of the population.

Avina Sarna, Lopamudra Ray Saraswati, Jerry Okal, James Matheka, Danmark Owuor, Roopal J. Singh, Nancy Reynolds, Sam Kalibala

https://doi.org/10.9745/GHSP-D-18-00241

“It’s Not Like Taking Chocolates”: Factors Influencing the Feasibility and Sustainability of Universal Test and Treat in Correctional Health Systems in Zambia and South Africa

Universal test and treat may be feasible even in highly resource-constrained correctional facilities. Sustainability and impact of such services require a supportive policy environment, robust service delivery systems, adequate resourcing, and close attention to the psychosocial factors influencing incarcerated persons’ willingness to engage in HIV treatment.

Stephanie M. Topp, Candice M. Chetty-Makkan, Helene J. Smith, Lucy Chimoyi, Christopher J. Hoffmann, Katherine Fielding, Stewart E. Reid, Abraham J. Olivier, Harry Hausler, Michael E. Herce, Salome Charalambous

https://doi.org/10.9745/GHSP-D-19-00051

Alternative Ready-To-Use Therapeutic Food Yields Less Recovery Than the Standard for Treating Acute Malnutrition in Children From Ghana

In Ghana, an alternative ready-to-use food (RUTF) formulation that met all specifications was not as good as standard RUTF in affecting recovery from acute malnutrition among children aged 6 to 59 months.

Kristin Kohlmann, Meghan Callaghan-Gillespie, Julia M. Gauglitz, Matilda Steiner-Asiedu, Kwesi Saalia, Carly Edwards, Mark J. Manarya

https://doi.org/10.9745/GHSP-D-19-00004

Evaluating WHO-Recommended Interventions for Preterm Birth: A Mathematical Model of the Potential Reduction of Preterm Mortality in Sub-Saharan Africa

Using the Maternal and Neonatal Directed Assessment of Technology (MANDATE) model, we estimate that WHO-recommended interventions could have saved nearly 300,000 lives in 2015. Combined interventions had the greatest impact. MANDATE can allow health officials to prioritize implementation strategies.

Jennifer B. Griffin, Alan H. Jobe, Doris Rouse, Elizabeth M. McClure, Robert L. Goldenberg, Beena D. Kamath-Rayne

https://doi.org/10.9745/GHSP-D-18-00402
Introduction of Community-Based Provision of Subcutaneous Depot Medroxyprogesterone Acetate (DMPA-SC) in Benin: Programmatic Results

Lay community health workers and facility-based health care providers in Benin were trained to administer DMPA-SC safely and effectively in 10 health zones. Community-based DMPA-SC was popular, particularly among new users of contraception, and could help the country achieve its family planning goals.

Tishina Okegbe, Jean Affo, Florence Djihoun, Alexis Zannou, Odilon Hounyo, Gaston Ahounou, Karamatou Adegnika Bangbola, Nancy Harris

https://doi.org/10.9745/GHSP-D-19-00002

Are Procured Quantities of Implants Adequate and Appropriate? Modeling Procurement, Inventory, and Consumption of Contraceptive Implants During Rapid Uptake

Recent rapid increases in implant procurement have not resulted in system overstocks to date. We found no standard factor for relating inventory quantities to consumption rates. Rather, that relationship requires specific understanding of the country supply chain, inventory control parameters, and current and future demand.

Laila Akhlaghi, Alexis Heaton, Yasmin Chandani

https://doi.org/10.9745/GHSP-D-19-00017

The Challenges of Transition From Donor-Funded Programs: Results From a Theory-Driven Multi-Country Comparative Case Study of Programs in Eastern Europe and Central Asia Supported by the Global Fund

Transitioning from donor funding toward domestic financing for HIV and TB programs in Eastern Europe and Central Asia presents major challenges. It will require a substantial multipronged approach through well-planned collective and coordinated responses from global, bilateral, and national partners.

George Gotsadze, Ivdity Chikovani, Lela Sulaberidze, Tamar Gotsadze, Ketevan Goguadze, Nertila Tavanzhi

https://doi.org/10.9745/GHSP-D-18-00425

Efficacy of a Digital Health Tool on Contraceptive Ideation and Use in Nigeria: Results of a Cluster-Randomized Control Trial

A mobile digital health tool piloted in Kaduna City, Nigeria, was efficacious in promoting positive contraceptive attitudes and encouraging women to adopt a modern contraceptive method, thus showing potential for reducing unmet need in Nigeria.

Stella Babalola, Caitlin Loehr, Olamide Oyenubi, Akinsewa Akiode, Allison Mobley

https://doi.org/10.9745/GHSP-D-19-00066
Adding a Question About Method Switching to the Method Information Index Is a Better Predictor of Contraceptive Continuation

Adding the question “Were you told about the possibility of switching to another method if the method you selected was not suitable?” to the Method Information Index (MII) was associated with better contraceptive continuation. This MIIplus variable includes another domain of quality of care, and thus better reflects voluntary contraceptive use and continuation.

Aparna Jain, Kumudha Aruldas, Elizabeth Tobey, Arupendra Mozumdar, Rajib Acharya
https://doi.org/10.9745/GHSP-D-19-00028

Associations Between Practices and Behaviors at the Health Facility Level and Supply Chain Management for Antiretrovirals: Evidence from Cameroon, Namibia, and Swaziland

Using antiretrovirals (ARVs) as tracer products, we identified the following key practices that may affect supply chain management at the facility level: order verification, actions taken when stock is received, changes in prescription and dispensing due to ARV stock-out, actions to ensure patient adherence, and communication with other affiliated facilities and higher-level supply chain management. We propose a set of indicators to measure these practices.

Diana Bowser, Laura Krech, David Mabirizi, Angela Y. Chang, David Kapaon, Thomas Bossert
https://doi.org/10.9745/GHSP-D-19-00063

TECHNICAL NOTES

Planning for Outcomes (P^4O) Modeling Tool: Estimating the Impact of Changing the Proportion of Injectable Progestins in the Contraceptive Method Mix

The interactive deterministic online modeling tool P^4O allows users to estimate how changing the proportion of injectable progestins in the contraceptive method mix might affect HIV and maternal and child health outcomes. With careful consideration for women’s individual choices, policy makers and program planners may use country-specific results to help inform programming and policy decisions.

Elena Lebetkin, Xiaoming Gao, Douglas Taylor, Lauren Y. Maldonado, Abdulmumin Saad, Markus J. Steiner, Laneta J. Dorflinger, Kavita Nanda, Timothy D. Mastro
https://doi.org/10.9745/GHSP-D-19-00062
The Extent to Which Performance-Based Financing Programs’ Operations Manuals Reflect Rights-Based Principles: Implications for Family Planning Services

Rights principles should be prioritized and more clearly stated in performance-based financing (PBF) guidance and operational documents. Additional research, including development of validated measurement metrics, is needed to help PBF programs systematically align with rights-based approaches to health care including family planning.

Marie S. Cole, Victoria Boydell, Karen Hardee, Ben Bellows

https://doi.org/10.9745/GHSP-D-19-00007
Human-Centered Design and Sustainable Malaria Interventions

Michael Macdonald, a Thomas Putzer b

Human-centered design provides a method to adapt malaria control interventions to be more closely aligned with a family’s convenience, comfort, and personal lifestyle, enabling a broader and more sustained culture of access and use.

See related article by Kim.

For every complex problem there is an answer that is clear, simple, and wrong. - H.L. Mencken (paraphrased), Prejudices: Second Series (1920)

The article by Kim et al.,1 “Using a human-centered design approach to determine consumer preferences for long-lasting insecticidal nets in Ghana,” raises 2 important questions. First, can the public sector adapt long-lasting insecticidal net (LLIN) designs to the user’s point of view rather than purely technical and cost considerations? Second, is it possible to segment the market for mosquito nets such that people who can afford to pay can acquire the nets they want commercially, while the public sector focuses limited resources on those most in need who are unable to afford to purchase nets?

These questions challenge current strategies on 2 counts. First, large international tenders for mosquito nets are awarded on unit costs that meet the minimum physical standards. Second, “universal coverage” demands that all persons living in malaria endemic areas receive a free standard net.

While treated mosquito nets can have a significant impact on malaria illness and death,2 some estimate that, depending on the community being studied, as much as half of the households that receive nets do not use them for their intended purpose.3 Excessive heat and reduced airflow are often cited as barriers to LLIN use,4 with reduced air flow ranging from 55% to 71% in 1 study of 11 commercial nets.5 Imagine how much more impact LLINs could make if design changes to increase end-user compliance were addressed.

Kim et al. explored the barriers to LLIN ownership and use among middle-class Ghanaians through a human-centered design (HCD) process that moves beyond the traditional public health tools of focus group discussions, household surveys, and trials of improved practices:

The result was a rich mix of data and the identification of key consumer insights regarding middle-class Ghanaians’ perceptions of self, their behaviors and attitudes related to malaria prevention, and their use of LLINs.

The study found:

… in most accounts [free public-sector LLINs] were inconvenient, uncomfortable, and not aesthetically pleasing, thus they were undesirable to use.

Suggested changes to the standard LLINs included a more convenient way to hang the net, a more attractive silhouette, and a zipper for ease of entry and exit. Previous LLIN design work in Ghana included the addition of a solar-powered light and fan.6

HCD, the process described by Kim et al., is an iterative approach to generating solutions that are firmly rooted in people, developing empathy with the end-user, generating an abundance of ideas, building tangible prototypes, and iteratively co-creating with people again. Four lenses of user desirability, business viability, technical feasibility, and sustainability typically serve as a framework, focusing on optimizing for the user experience while ensuring viability at scale. HCD has become a key element in the Center for Innovation and Impact7 at the United States Agency for International Development and at the Design for Health Initiative.8

The value of HCD for improving LLIN access and use is well recognized. However, the business viability of implementing HCD solutions that go beyond addressing only the technical feasibility is challenged by the current system. Currently, manufacturers are simply “vendors” rather than “partners,” with a focus on manufacturing uniform LLINs, with minimum specifications at the

a Associate Editor, Global Health: Science and Practice Journal, and Consultant, Catonsville, MD, USA.
b SC Johnson & Son, Inc., Racine, WI, USA.
Correspondence to Michael Macdonald (macdonaldm@macito.net).

Global Health: Science and Practice 2019 | Volume 7 | Number 2

148
comfort, and personal lifestyle. Ultimately, our hope is that, as HCD emphasizes, people impacted by malaria will be enabled to guide us according to their priorities, how they live, and what is important to them. In turn, we will have a chance to create sustainable solutions that fit their lifestyles and will prevent malaria.

With the uncertainty of future global health funding, it is more important than ever that we optimize available resources to segment LLIN delivery strategies—that is, facilitate the growth of a consumer market for those who can afford to make purchases and concentrate public health resources on those who cannot, offering a diversity of LLIN products that fit the needs and preferences of each.

In his commentary on the survival of the Global Fund to Fight AIDS, Tuberculosis and Malaria, Richard Horton notes:

There are, of course, other questions the Fund must consider. How far should it embrace the private sector? What is its strategy for middle-income countries? How can the Fund leverage domestic investments...

In the context of LLIN distributions, this could be rephrased: How do we facilitate a commercial LLIN sector, what is the strategy for middle-income families, how can the Fund leverage consumer investments?

Kim et al. provide a step in that direction:

We have since shared our consumer insights and preliminary ideas for new design features with current manufacturers globally who supply LLINs in Ghana. We hope these partners take this information into consideration as they make decisions about current and future LLIN supply, demand, and marketing and will pursue pilot testing of new net designs for the private sector retail market in Ghana.

They conclude with a statement that seems to sometimes be lost on policy makers:

For a health technology such as the LLIN to produce a benefit, it has to be used.

There is much to be learned from consumer product development and HCD to solve complex problems for which we thought, with our hubris, we already knew the simple solution.

Competing Interests: None declared.

REFERENCES


Supervision of Task-Shared Mental Health Care in Low-Resource Settings: A Commentary on Programmatic Experience

Christopher G. Kemp, a Inge Petersen, b Arvin Bhana, b,c Deepa Rao a,d

Task-shared mental health care programs in low-resource settings often incorporate supervisory structures that would be difficult to implement at scale, and many rely on foreign specialist experts as supervisors. Future programs could leverage peer supervision, technology, competency assessments/fidelity checklists, and other tools. Mental health care specialists will require training, support, and incentives to supervise generalist care providers.

INTRODUCTION

Mental disorders are the leading cause of years lived with disability globally. Yet in low- and middle-income countries (LMICs) and other low-resource settings, 75% of people in need of treatment for mental disorders never receive care. Effective services that are feasible, scalable, and sustainable in the context of critical shortages of financial and human resources are needed to bridge this treatment gap. Few health systems in LMICs can rely exclusively on specialists to deliver mental health interventions, nor can they afford to develop mental health programs in parallel to other services. Instead, they have to rely on existing cadres of health care workers and constrained financial resources to expand access for mental health services.

One promising approach has been to deliver psychosocial or pharmacological services via task sharing. Task sharing is an arrangement in which generalists—non-specialist health professionals, lay workers, affected individuals, or informal caregivers—receive training and appropriate supervision by mental health specialists and screen for or diagnose mental disorders and treat or monitor people affected by them. Systematic reviews of task-shared mental health services in low-resource settings have demonstrated that the approach can be acceptable and feasible and can lead to substantial improvements in patient health outcomes, even in settings with few available specialists.

To ensure generalist providers adopt evidence-based mental health services and deliver them with fidelity, task-sharing programs must incorporate effective systems for ongoing training, supervision, and mentorship. Initial trainings are a necessary but insufficient step toward building the confidence and competence of mental health clinicians. Supervision and mentorship are essential to developing the feedback loops that correct negative behaviors and reinforce positive behaviors as part of the cycle of learning, doing, and reflecting. In high-resource settings, supervision has been widely recognized as key to the development of a clinician’s skills. Programs without ongoing supervision have been found to have low intervention fidelity and clinician competency, and established programs without supervisory support can risk significant declines in service delivery within 2 years. Indeed, the type of training received has been shown to be less important than the dosage of supervision in predicting clinician adoption of and fidelity to evidence-based mental health services. Supervised mental health clinicians receive essential emotional support and are less likely to experience burnout.

The relative effectiveness of different supervisory models for task-shared mental health services in low-resource settings remains understudied, although recent calls for research suggest that a change is imminent. Little is known about the range of supervisory models already developed and implemented as part of task-shared mental health care in low-resource settings. An exploration of these models would offer support to future programs as staff plan, design, and implement task-shared programs. Our objectives were to provide an overview...
of the literature on the supervision of frontline and mental health care workers in low-resource settings, to describe and draw lessons from the experiences of implementers of task-shared mental health services in these settings, and to offer evaluative commentary for consideration by future investigators and implementers.

**OVERVIEW OF SUPERVISION MODELS**

Supervision of frontline health care workers—including but not limited to those delivering mental health care—may take many forms. Most broadly, supervision refers to the cyclical process in which a senior professional or team sets expectations for the practice of health care workers at a lower level in the health system, observes and/or audits that practice, assesses whether it meets expectations, and provides guidance or takes corrective action. Supervisors employ a wide range of activities to carry out these functions, and health systems may focus on and prioritize some supervisory functions over others. Depending on that focus, models for supervision fall along a spectrum of 3 general categories: traditional supervision, supportive supervision, or mentorship.

Traditional supervision is distinguished by its focus on oversight for the purpose of identifying problems, with little emphasis on guidance or support. Many LMIC health systems can trace the use of this form of supervision to colonial powers that enforced rigid hierarchies with the goal of ensuring compliance among local staff and lower-level workers. Traditional supervision imposes the needs of the health system onto providers, while often failing to address the needs of providers or the needs of the patient population. The premise is that satisfactory performance can only be achieved through close control and punitive measures. Problem solving is reactive and episodic, rather than proactive and continuous. Under this model, frontline health care workers are not empowered to identify and solve issues on their own. Traditional supervision typically occurs through short visits by external supervisors to health care facilities and the completion of routine forms and checklists.

In contrast, supportive supervision has been described as a process of strengthening relationships within the health system, with a focus on continuously identifying and resolving problems, optimizing the allocation of resources, and promoting teamwork and open communication, all with the goal of improving quality across all levels of the health system. Supportive supervision incorporates self-assessment and assessment by peers, as well as community input, shifting the focus of supervision from the supervisor-supervisee dyad to the entire workforce. It uses a practical system of objective measures to foster improvements in the procedures, personal interactions, and management in primary care facilities. Supportive supervision has been shown to be conducive to improvements in health worker performance and to a more general strengthening of health systems in low-resource settings. In sub-Saharan Africa, evidence suggests that supportive supervision increases job satisfaction and worker motivation. Supportive supervision is considered best practice and is the model specifically recommended in the context of task sharing: World Health Organization (WHO) guidelines state that supportive supervision should be regularly provided to all task-shared health workers, and that supervisors should themselves be competent and have appropriate supervisory skills.

At the far end of the supervision spectrum, mentorship broadens the focus from building skills related to a specific intervention or health issue to fostering the development of a learner’s professional career. Mentorship has been found to be an effective intervention for knowledge translation. Cultural congruency between mentor and mentee may be important to ensuring an effective mentoring relationship.

In addition to the general categories outlined above, supervision of frontline health care workers—including mental health care workers—can vary by dosage, mode or level, tools used, and the amount of supervisor training. Limited evidence of the comparative effectiveness of the various combinations of each of these variables exists; even the effectiveness of supervision compared to no supervision remains unclear and understudied. No ideal dosage of supervision has been identified—whether weekly, monthly, yearly in frequency, or 1 hour, 1 day, or 1 week in duration—although evidence does suggest that the quality of supervision is more important than its frequency. Peer supervision has been found
Few in-depth descriptions of models of supervision exist for task-shared mental health care in low-resource settings. We sought to describe and learn from models for supervision that have been developed for task-shared mental health services in low-resource settings.

Specific to task-shared mental health care in low-resource settings, the literature offers few in-depth descriptions of models of supervision and even less evidence of comparative effectiveness. One early report, published in 2001, contains brief overviews of supervisory structures for pharmacological mental health care services integrated with primary care in Guinea-Bissau, India, Iran, and Nicaragua, and emphasizes the importance of ongoing supervision beyond the period of training.40 A series of case studies from Latin America and the Caribbean and sub-Saharan Africa, describing the implementation of both psychosocial and pharmacological task-shared services, specifically highlights the role of supervision in helping generalists who report feeling depressed or stressed when they start delivering mental health care interventions. These case studies also suggest that a minimum number of mental health specialists are required to supervise generalists and to provide specialized referral treatment services.41,42 Systematic reviews and cross-country studies of task-shared psychosocial and pharmacological services reiterate the need for ongoing supervision at the community and primary care levels to help generalists overcome difficulties and strengthen their capacity to meet patient needs, although descriptions of supervision models are limited.43–45 A recent systematic review of LMIC-based trials of task-shared psychosocial treatments for common mental disorders found that just over half of the studies in the sample described the format, method, frequency, or cadre of supervisor used.10 Of those that did report these details, all conducted supervision in person; group-based supervision was more common than individual supervision; most used individual cases to guide supervision; most used expert specialists as supervisors; and most conducted supervision weekly.10

One well-documented approach to task-shared mental health supervision—focused specifically on psychosocial treatments—is the apprenticeship model: a collection of training and layered supervision methods originating with researchers at Johns Hopkins University, named after the model used by many crafts and trades.46 It is distinguished by its inclusion of 3 types of individuals: counselors, supervisors, and trainers.47 Counselors may be any type of mental health service provider, including community members trained to deliver a psychosocial intervention, while supervisors are counselors with the expertise or skills necessary to support other counselors. Trainers are experts from outside the service delivery context. The apprenticeship model has 5 steps: (1) selection of counselors and supervisors, (2) training, (3) practice groups, (4) supervised expansion of skills, and (5) mutual problem solving. It is relatively time and resource intensive: new counselors may shadow and observe supervisors for several weeks or months, after which supervisors perform direct clinical observation of counselor or service delivery, incorporating periods of reflection and debriefing. Trainers have tended to be specialists or academics from high-resource settings.48 Nonetheless, investigators have successfully applied the apprenticeship model across numerous settings, including the Democratic Republic of the Congo, Haiti, Iraq, Thailand, and Zambia.48–53

In summary, supervision of frontline and mental health care workers falls along a spectrum from traditional supervision to supportive supervision and to mentorship. Supervisory models may vary by dosage, mode or level, tools used, and the amount of supervisor training. Limited evidence of the comparative effectiveness of these models exists. Even less evidence exists that is specific to the supervision of task-shared mental health services in low-resource settings.

**EXPERIENCES FROM THE FIELD**

In an effort to further describe and learn from models for supervision that have been developed for task-shared mental health services in low-resource settings, we interviewed key informants, including researchers, program managers, and clinicians. We identified potential informants through the literature and through our professional networks, based on their experience designing, supporting, supervising, or delivering task-shared services. We used a snowball recruitment method to increase variation in position...
type, funder, project type, and supervisory model. We contacted potential informants by email. The first author conducted all interviews over the phone, over Skype, or in person, using a semi-structured interview guide with open-ended questions. Interviews lasted up to 1 hour and covered informants’ experiences with task-shared mental health interventions, the models for supervision that they developed or used, and their thoughts about how to structure future models for supervision. We digitally audio-recorded and transcribed all interviews. The first author used line by line coding to identify themes according to an inductive approach of constant comparison and content analysis. The Human Subjects Division of the University of Washington determined that this study qualified for exemption status under 45 CFR 46.101 (b).

We contacted 21 potential informants; 5 refused or did not respond. We interviewed 16 informants between October 2015 and January 2017. Most were researchers, and most worked in sub-Saharan Africa. Over half were from an LMIC. Informants described models of supervision from a range of different research projects or service delivery programs. Most projects and programs were based in sub-Saharan Africa, and most were funded as short-term studies by research institutes or bilateral donors. Although most offered psychosocial interventions, several incorporated pharmacological treatment. Almost all were implemented as clinical trials, led by public mental health experts based in high-resource settings, with limited funding or capacity for long-term sustainability of service delivery.

Informant experiences reflected 5 broad themes: movement from research to scale-up; building capacity for supervision by specialists; social hierarchies and supportive supervision; technological opportunities; and allowing for context, fluidity, and heterogeneity. We describe each of these below.

Movement From Research to Scale-Up
Most of the models for supervision were designed by researchers to meet funder requirements and promote the fidelity of trial interventions. These models were often not formally structured, manualized, or designed to be sustainable or implemented at scale. Almost all informants discussed using specialists from research institutions in high-income countries, or high-resource settings, as primary or secondary supervisors, supporting service delivery through occasional calls or on-site visits. This approach has clear implications for the long-term feasibility of the programs and may limit the external validity of trial results.

When it comes to mentoring and supervision of [counselors] over the long term, things fall apart to a certain degree because either the pilot studies that have been done they’re supervised by the [principal investigator] or they are temporary. [Male Academic]

However, some programs were moving to more scalable and sustainable models of supervision, driven in part by a shift from acceptability, feasibility, and effectiveness trials, towards implementation and scale-up research. Informants identified a clear need for the dissemination of evidence-based models for supervision that were structured, feasible, and sustainable. Commonly cited tools for this purpose included manuals, competency measures, fidelity checklists, and decision trees. Several informants described using manuals to describe and standardize supervision across supervisors and project sites and using fidelity checklists to monitor generalist practice.

Building Capacity for Supervision by Specialists
Mental health specialists are not generally trained or paid to supervise generalists. They do not have the time or skills necessary to be supportive and offer mentorship. This limitation is especially true in low-resource settings, where specialist time is a precious resource. Specialists are trained as clinicians, not managers, yet task sharing may require them to facilitate and guide the work of dozens of generalists. Informants suggested that specialists need additional training in supervision and personnel management to manage teams of task-sharing mental health workers.

What’s very clear is, in fact, we can’t just expect... people who’ve been trained in clinical psychology to be able to do this. It needs to be built into the core competencies for the preservice training and then they, themselves, need supervision to be able to provide supervision. It’s actually caused a bit of a problem for us because the counselors are not—as far as I’m concerned—are not receiving adequate supervision because they come to the office once a week where [Program Manager] gives them supervision but it’s not working very well. [Female Academic]

Although some specialists may be willing and eager to be trained to support the expansion of mental health services via task sharing, informants suggested that many are not. The specialists need additional training in supervision and personnel management to manage teams of task-sharing mental health workers.
are often already overwhelmed with their service delivery duties, especially given the profound scarcity of human resources, and are therefore unable to add the responsibilities of generalist supervision. Some informants also reported that specialists may feel protective of their skillset and of the services they provide, acquired after years of training, and resist the notion that generalists can task share these services safely and effectively. 

One of the issues that’s come up a lot in other interviews has been of special[ist] staff who are trained as service providers in psychology or psychiatry who are not comfortable diversifying their role to include supervision of task-shared health workers. [Male Academic]

Social Hierarchies and Supportive Supervision
Informants frequently discussed the fact that task sharing unavoidably creates or reinforces a power imbalance across the specialist and generalist hierarchies. Specialists in low-resource settings, especially psychiatrists, tend to be of high socioeconomic status and male. Generalists, such as nurses and community health workers, however, tend to be female and may be from lower socioeconomic status.

Not only that, but you have psychologists who are trained in urban capitals, upper middle class, they don’t want to work with community health workers, they speak a different language. [Male Academic]

Informants indicated that such imbalances make it difficult for generalists to speak up during individual or group-based interactions with supervisors, making such interactions unproductive. Several informants reported the successful use of peer supervision as an alternative model that circumvented imbalances in power.

I think the most important strength was that it encouraged the peers to really reflect and to be more alert and to contribute because sometimes what also happens is what we had noticed... is that the peers often feel intimidated. You know, they feel that when the supervisor and the expert supervisor is present in the group it’s almost as though they don’t feel very comfortable voicing their opinions because I guess they recognize that they’re trained just specifically in one thing and they don’t feel confident enough... [Female Program Manager]

Several informants had assumed that specialists or program managers would be able to visit generalists for on-site supervision at least once per month, only to find that such travel was not feasible. Consequently, some projects resorted to supervision at a distance, via telephone, text messages, or the Internet. Several projects found WhatsApp groups to be an effective, affordable tool for peer-to-peer, group-based supervision and discussion.

We had a WhatsApp network with all the facilitators... They might get a reply from another colleague or other peer, but... the messages were supervised or overseen, perhaps, is a better word, by a supervisor that whenever that person felt they needed to intervene, then it will come into the network with a message for everybody. [Male Academic]

One project used tablets to monitor implementation and service delivery and to send advice and feedback to generalists. Many successfully used video or audio recording to supervise patient encounters or therapies, sharing examples during group supervision for peer-based review and critique. Although such recording was often necessary to promote the fidelity of interventions delivered as part of research-based trials, informants suggested that it could be scaled as part of routine service delivery and supervision.

Now that possibility would be... for the counselors to record their sessions and for those supervisors—if they’re not able to go to the sites because of distances and things like that—to record the sessions and then try to do some supervision over the phone, having... listened to the recordings. [Female Academic]

Allocating for Context, Fluidity, and Heterogeneity
All informants emphasized that models for supervision must be adapted to the specific resources and needs of a setting. What works in one district may not work in another, and the model for supervision applied to one package of services may not translate to another package.

I think it’s context specific because there are characteristics of the context that have to be taken into consideration. I mean Ministry health infrastructure, existing staffing, rural or urban, how it’s organized geographically, what’s the condition of the community health workers, what’s the quality of the health care system itself. [Male Academic]

Technological Opportunities
Several projects were implemented over large areas, often with poor roads or difficult terrain.
Moreover, informants suggested that supervisory structures have changed and will continue to change over time to accommodate the development of services and the shifting needs of caregivers.

It’s something that evolved over time. We tried quite a lot of things. We tried getting doctors to drive to the clinic and on a weekly basis work there for referrals and that just didn’t work. We tried to get cases—other, severe cases—to go immediately to the tertiary facility and then that wasn’t very reliable. We then realized that the best thing we could do is empower the lay health workers as much as possible. [Male Academic]

In summary, we identified 5 broad themes. Our informants suggested that task-sharing programs are shifting from effectiveness research to scale-up; few specialists in low-resource settings have the time or skills necessary to supervise generalists; task sharing may create a power imbalance between specialists and generalists; technological solutions may improve the practicality and reduce the cost of supervision; and finally, models for supervision should be adapted to each context and will need to change over time.

### DISCUSSION

Our exploration of supervisory models for task-shared mental health care suggests that a shift is underway as programs move from effectiveness research to implementation and scale-up. Initially, investigators designed models for supervision to meet the needs of clinical trials. These benefited from the extra financial and human resources available for research, limiting their external validity, and were rarely standardized. None have been evaluated independently. As we implement programs at scale, however, it will be critical to incorporate models for supervision that are carefully planned and evidence based. We can use standardized manuals, routine fidelity checklists, and other evaluative components to promote effective and sustainable supervision of task-shared mental health services.

The themes documented above align with previous studies emphasizing that task sharing requires adequate approaches to and resources for training, supervision, and emotional support. Training generalist staff to provide mental health care without considering their workloads or providing ongoing supervision can lead to inappropriate treatment. Without support, task sharing can become task dumping as generalists are overloaded with tasks they cannot perform. Where available, specialist mental health workers can be used to supervise mental health care in the primary care and community settings, although their supervisory roles must be clearly delineated and compensated. Sufficient numbers of specialist mental health human resources are required to provide effective and sustained supervision and support to generalists. Specialists should also be adequately prepared to supervise generalists through appropriate preservice training in supervision and mentorship—this aspect will require a reform in mental health specialist core competences and the engagement of specialist communities and other relevant stakeholders. In high-income countries, it is increasingly common for psychological training programs to offer courses in supervision, teaching specialist trainees the skills necessary to provide clear feedback and structured assessment of clinical practice. The American Psychological Association now considers clinical supervision to be a core competency of the health service psychologist, and it publishes guidelines for clinical supervision. In low-resource settings, public mental health services may also need to delineate generalist supervision as part of the specialist job description and establish achievable targets for the number of generalists supervised by each specialist. This approach would help specialists to have protected time for supervision and motivate them to meet their targets.

Use of established models, tools, and technology may improve the scalability and rigor of supervision, if it is adapted to context and used appropriately. The apprenticeship model offers a tried and tested method for training, mentoring, and supervising task-sharing generalists in low-resource settings, and it may improve clinical competency and provider confidence—although the reliance on expert trainers may be a significant barrier to scale-up. To ensure feasibility at scale, locally available competent trainers are required. To date, trainers of programs using the apprenticeship model have often been specialists from universities in high-income countries. In the context of limited specialists in LMICs, peer-based supervision using standardized measures of task-shared provider competence may be a highly acceptable and valid complement to expert supervision, and it would help mitigate the bottleneck of limited specialists in these contexts. WhatsApp groups, telephone, Skype, and other technologies provide valuable spaces for individual-, peer-, and group-based supervision at a distance and on demand, improving feasibility in areas where regular in-person meetings are challenging and encouraging.
ongoing peer-to-peer learning. A qualitative study of supervision of community health workers using WhatsApp groups in Kenya suggested that the groups spur healthy competition and team building. However, implementers may need to be cautious with this approach and consider how best to monitor and engage in groups to ensure discussions are not misleading, distracting, or harmful. Implementers of a task-shared collaborative care model for psychotherapy in rural Nepal have shown that regular telephone calls can provide a valuable supplement to less frequent in-person supervision by a specialist.

Moving forward, it will be essential for task-shared mental health services to incorporate models of supervision that are feasible at scale and adapted to local contexts and resource levels. The Table outlines possible models for supervision according to different resource levels, expanding on a hierarchy proposed in the WHO Mental Health Gap Action Programme Operations Manual. Although we advocate for increased global investment in mental health, current approaches should be tailored to be scalable and sustainable given weak and under-resourced health systems. It will not be feasible for large programs to rely on academic researchers or experts from high-resource settings for ongoing supervision or training. Rather, programs should maximize use of in-country staff. Although the most effective models of supervision may also be the most time and resource intensive, options are available for programs with fewer resources. Supervision can be layered, such that experienced peers provide the bulk of supervision and scarce specialists are reserved for special cases or infrequent supervisory sessions. In many settings, it will be not be practical for supervision to occur in person. In such settings, programs could consider audio- or video-recording trainee service delivery to check fidelity, telephone or Skype calls for individual supervision, and WhatsApp or other group texts for group or peer supervision. Finally, manualized and routine supervision—supported by tools like fidelity checklists to monitor and assess provider competence—is likely to yield superior outcomes in scarce resource contexts, compared to ad hoc drop-in supervision.

CONCLUSION

Supervision is an understudied but critical component of task-shared mental health programs in low-resource settings. As interventions move from development to implementation and scale-up, models for supervision that are feasible for dissemination are increasingly being developed. In the absence of adequate numbers of specialists to provide supervision, technological solutions like audio recording and WhatsApp groups supported by supervisor guides and fidelity checklists can help promote better quality supervision as well as contact with supervisees. Further research is necessary to evaluate models for supervision across different programs and contexts.

Acknowledgments: We would like to acknowledge and thank our key informants. We also thank Dr. Milton Wainberg (Columbia University) and Dr. Bradley Wagenaar (University of Washington) for valuable input on plans for data collection.

Funding: CGK was supported by grant number F31MH112397 from the National Institute of Mental Health.

Competing Interests: None declared.

REFERENCES


Peer Reviewed

Received: September 6, 2018; Accepted: April 2, 2019


© Kemp 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-18-00337
Using a Human-Centered Design Approach to Determine Consumer Preferences for Long-Lasting Insecticidal Nets in Ghana

Sharon Kim, a* Danielle Piccinini, b* Elorm Mensah, c Matthew Lynch b

Through focus group discussions and human-centered design exercises, middle-class Ghanaians communicated the need to address convenience, comfort, and aesthetics when designing a bed net for their demographic. Illustrative attributes for consideration by private-sector manufacturers include a more convenient way to hang the net, a more attractive silhouette, and a zipper to provide ease of entry and exit while keeping the area sealed from mosquitos.

ABSTRACT

Background: A human-centered design approach, paired with traditional research methods, was used to explore consumer preferences of middle-class Ghanaians for a long-lasting insecticidal net (LLIN) to be designed for the private-sector retail market.

Methods: In March 2017, we conducted 9 focus groups with urban and rural middle-class Ghanaians across Ashanti, Greater Accra, and Western regions. A total of 78 participants (51 adults and 27 boarding school students) were involved in the focus groups. Participants were asked for their input on topics related to malaria prevention, LLIN perceptions and use behavior, as well as general consumer preferences related to the home and bedroom. They participated in a variety of exercises, such as rank ordering their preferences of various accessories that might be bundled with an LLIN and interacting with actual LLINs of different sizes and designs. The data were gathered and analyzed, using micro-interlocutor analysis framework to capture emergent themes.

Results: LLINs are currently available through free distribution channels, but in most accounts, participants reported that the nets were inconvenient, uncomfortable, and not aesthetically pleasing, thus they were undesirable to use. For example, several participants described the process of hanging as well as entering and exiting the LLIN as challenging, stressful, and/or tedious. In addition, use of LLINs was considered to make people feel even hotter in an already warm climate as well as to leave users feeling confined within a small space. Finally, many participants discussed how to improve the look of LLINs including suggestions for additional colors, shapes, and hanging mechanisms to make the LLINs more compatible with their desired bedroom and home décor. Based on the participants’ responses, we concluded that they would prefer LLINs that better reflect contemporary sensibilities. We devised and tested different LLIN attributes to address these points, focusing on a more convenient way to hang the net, a more attractive silhouette, and a zipper that allows the user to enter and exit with ease while still maintaining a sealed, mosquito-free space. A separate discrete choice experiment confirmed the attractiveness of these attributes by capturing the target population’s willingness to pay for LLINs with said preference-congruent attributes.

Conclusion: Our human-centered design approach yielded consumer insights and preferences for novel LLIN designs for the private-sector retail market in Ghana. If this net design is successful, it could increase LLIN use among the middle class and catalyze the purchase of LLINs to support ongoing malaria control efforts.

BACKGROUND

The long-lasting insecticidal net (LLIN) is the mainstay of vector control for malaria prevention. Rapidly scaled-up dissemination through mass distribution of free LLINs has been credited with 68% of the 663 million clinical cases of malaria averted between 2000 and 2015.1 While the impact of widespread LLIN ownership and use has been substantial, a recent study of the LLIN use-access ratio, an indicator of usage given access to LLINs, has provided an additional important dimension to consider with regard to malaria prevention. Comparison of use-access ratios in Ghana show that among people who own LLINs, those in wealthier urban households have significantly lower LLIN use compared with those in rural households, particularly rural households in lower wealth quintiles.2
The reasons for inconsistent or non-use of LLINs vary and are not always fully known or understood. Understanding the barriers to LLIN ownership and use is an important factor in the allocation of resources for LLIN distribution by the National Malaria Control Program. If, for instance, the private sector can provide the middle class with affordable LLINs that meet consumer preferences, the National Malaria Control Program might be able to adjust its distribution efforts to allocate a greater number of free LLINs to populations at higher risk for malaria that use LLINs at higher rates.

In our work, we focus specifically on the LLIN usage of the middle class in urban and rural areas of Ghana. This demographic is growing and draws significant interest from public health advocates and researchers, among others, during the current period of unprecedented economic growth in Ghana. Our research is part of a larger project that has the overarching purpose of catalyzing a viable commercial market for LLINs. To help achieve this goal, we focused on the sector of the population that can afford to purchase an LLIN, the Ghanaian middle class. Understanding some of the reasons for lower use of LLINs despite access within this population might also help reveal promising ways to reduce the gap (e.g., via LLIN modification, LLIN diversification, distribution channel, LLIN promotions, social behavioral communication change efforts).

In Ghana, mass campaigns typically distribute LLINs once every 3 years and are designed to achieve universal coverage by allocating 1 LLIN for every 2 people per household. During the periods between mass campaigns, LLINs are available through continuous distribution channels (e.g., certain services at health facilities) for biologically vulnerable populations, such as pregnant women when they attend their first antenatal visit and infants when they receive immunization for measles. LLINs are also provided through schools to students in grades P2 and P6. Should households not receive enough LLINs through these distribution channels, their options for obtaining an LLIN through other means are extremely limited. In other words, Ghana has no private-sector retail market for LLINs to speak of. According to a retail audit commissioned by the Private Sector Malaria Prevention Project, in 2017 only 7% of retailers in our 3 areas of focus (i.e., Ashanti, Greater Accra, and Western regions) sold bed nets. In addition, 95% of the total number of bed nets found on the market were claimed to be insecticide-treated, although the legitimacy of such statements was unverified and some of our interviewees questioned it. Many middle-class households in our areas of interest (approximately 82%) are using other malaria prevention methods such as aerosol insecticide sprays and coils.

The LLINs that are available through public-sector distribution channels may vary in terms of shape, size, color, and material. In Ghana, these LLINs are commonly rectangular in shape and made to fit a double-size bed (approximately 190 cm × 180 cm × 150 cm). They are generally standard shades of white or blue in color and are made of insecticide-treated polyester or polyethylene. Notably, although there is some modest variety in terms of LLIN features, people do not get to choose what type of net they receive in mass campaigns or through continuous distribution channels. Upon receiving an LLIN, people are advised to hang it outdoors in the shade for at least 24 hours prior to using it. Once it has been aired out, the LLIN must be hung such that it covers the entire sleeping area. People usually tie strings to the loops in the top four corners of the net and connect them to nails or hooks in the ceiling, wall, or other supports. The standard rectangular LLIN is close-sided and must be fully tucked under the mattress or mat to create a sealed environment for effective vector control (Figure). Typically, users tuck the net most of the way around the mattress, leaving a space to crawl into bed, and then tuck in the open section from inside the net.

Through a market analysis, we determined that LLINs designed for retail sale in Ghana would need to be noticeably different from the free LLINs distributed through public-sector channels, in ways that are attractive to target consumers. First, we found, unsurprisingly, that consumers are not likely to buy a product that is identical to one they can get for free elsewhere. Second, our target consumers were willing to pay for a differentiated LLIN provided it met their preferences. Our operating premise was that middle-class consumers will be more likely to use an LLIN that fits their stated consumer preferences and that regular use will protect users from malaria infection and reduce transmission rates by helping to control the mosquito population. Although attractive LLIN design options may be important for all users, our first goal was to help catalyze a retail market for LLINs among a population that can financially support it. Our work could be extended in the future to reach other users who were not represented in this single study.

We wish to emphasize that our project pertained to LLINs only. Our research showed that a
desire for malaria protection was a driver for purchasing prevention tools, and malaria prevention is very much a concern of middle-class Ghanaians. The responses indicated that many households consider the LLIN a desirable good but that it could be made more desirable. We feel a sufficient pool of demand exists in the middle class to drive the reestablishment of an LLIN market in Ghana. We expect that once this market is strengthened, it will expand to meet demand from other consumer segments of society over time. It should also be noted that the rationale behind these project efforts was not expected impact on disease burden among people in this segment of society (which would likely be small, given their lower risk), but rather building and expanding a sustainable market for LLINs that fully recovers costs.

Similar to other malaria prevention experts who have commented on the potential for redesign, we acknowledge the opportunity to put forth new LLIN designs to meet the contemporary needs of individuals living in malaria endemic areas. Thus, we sought to create new LLIN designs to meet one of our primary project goals of facilitating a functioning and competitive retail market for LLINs in Ghana. Our project focuses on the middle class because in order to catalyze a commercial market for LLINs, we first need to target the population with the economic means to purchase such a product. An in-depth exploration of the middle-class Ghanaian context was necessary in order to create LLIN designs that would meet that population’s specific needs. To that end, we took a human-centered design approach to our market research and product development efforts.

### METHODS

**Setting**

The bulk of this work was conducted in Ghana in the Ashanti, Greater Accra, and Western regions, with some work conducted in Baltimore, Maryland (USA). These areas of Ghana are ideal locations to sell LLINs because malaria is endemic and there are proportionately lower levels of poverty. In other words, a significant number of people in these areas are at risk of malaria and can afford to buy commercial LLINs.

**Human-Centered Design Process**

Human-centered design (HCD) is an approach to creative problem solving that prioritizes direct engagement with various stakeholders to glean insights that may be critical to designing products that are both novel and useful to a given market or audience; HCD also endorses iterative prototyping and testing. HCD is considered an umbrella...
category of various design approaches (e.g., user experience design, design thinking) and is a well-established part of certain business practices, such as product design and development. More recently it has begun to take hold in other professional sectors, including public health (e.g., see the Design for Health Initiative by the U.S. Agency for International Development and the Bill & Melinda Gates Foundation).11 Although inherently dynamic and iterative, the HCD process is often described in a series of simple steps (Table 1).12 We provide these steps as a reference, highlighting that it is a common way to frame this type of work; however, it is important to note that the steps are not executed as cleanly in reality as the table might suggest. For instance, the Empathize work, when the team immerses itself in the context of the key stakeholders, not only happens at the beginning but ideally is present throughout the other steps as both a guiding principle and a “check and balance” of human-centeredness.

Other user-centered approaches such as Trials of Improved Practices have been featured in selected public health projects, for instance, in efforts to improve nutrition planning13; however, we specifically chose to apply HCD because it allowed us the greatest flexibility in our transdisciplinary approach to working with our key stakeholders. Further, it is a practice that was familiar to our partners in both the private and public sectors. We used HCD in conjunction with traditional methods of qualitative and quantitative research (e.g., household surveys and focus group discussions [FGDs]). The purpose of this mixed-methods approach was not only to capture existing consumer preferences but also to tap into or discern latent preferences for a consumer product that does not yet exist. In other words, we needed to have a solid understanding of the consumer psychology within this population in order to create new LLIN designs that would be desirable to it. For this purpose, HCD was especially helpful because it encouraged individual participants to express thoughts and feelings about their experiences related to LLINs and malaria that would be more difficult to capture in other ways. For example, exercises and activities were designed to allow participants to brainstorm, imagine, and share intimate details regarding their domestic behaviors and routines. The result was a rich mix of data and the identification of key consumer insights regarding middle-class Ghanaians’ perceptions of self, their behaviors and attitudes related to malaria prevention, and their use of LLINs.

**Literature Review**

The Empathize phase of our HCD work began with an orientation to our target population and their current malaria prevention practices. We conducted secondary research including a targeted scan of peer-reviewed and gray literature related to malaria prevention, LLINs, and associated development and public health initiatives in Ghana. To further contextualize this research, we conducted the aforementioned market analysis as well as a series of supplementary interviews with key stakeholders. Our explicit goal was to gain a deeper understanding of the contemporary attitudes regarding these topics as well as the human experiences related to them, which could not be gleaned directly from the literature review.

**Market Analysis**

The market analysis included both qualitative (key informant interviews with different members of the supply chain) and quantitative (household surveys, discrete choice experiments, and retail audits) work. Through a combination of key informant interviews with 12 members of the LLIN supply chain and 271 retail audits, we sought to gain a better understanding of any reluctance to invest in an LLIN retail market in Ghana. The most commonly mentioned reason for a lack of interest was weak consumer demand. Owing to this low demand, manufacturers tended to focus their sales efforts on large institutional buyers like the

<table>
<thead>
<tr>
<th>TABLE 1. Steps of Human-Centered Design Process</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Empathize</strong></td>
</tr>
<tr>
<td>• Conduct secondary research</td>
</tr>
<tr>
<td>• Identify key stakeholders</td>
</tr>
<tr>
<td>• Engage with them in various ways to understand their circumstances and psychology</td>
</tr>
</tbody>
</table>

"We sought to gain a better understanding of any reluctance to invest in an LLIN retail market in Ghana."

A solid understanding of the consumer psychology within the study population was needed to design a desirable LLIN.

Global Health: Science and Practice 2019 | Volume 7 | Number 2

163
Ministry of Health and NGOs, whereas distributors and retailers put their efforts toward importing, stocking, and selling fast-moving consumer goods, which do include some mosquito control products (e.g., insecticide sprays and coils). Manufacturers and distributors were also specifically concerned about competition from counterfeit LLINs and LLINs leaked from free distribution channels. Despite these concerns, LLIN supply chain personnel were generally excited about the opportunity to catalyze a new market.

**User Interviews**
For our supplementary interviews with key stakeholders, we asked our subjects, including local community experts, consumer product managers, and malaria project team members (based in Ghana and the United States), to describe their own experiences with LLINs and why they did or did not use them. For example, one person described not using an LLIN because she had easy access to antimalarial drugs. Another described how the fear of scorpions and other vermin motivated him to use an LLIN regularly to get a good night’s rest. Although this information was not necessarily generalizable to a wider population, the interviews helped us unpack the psychology of an LLIN user of some financial means. In capturing and reviewing their responses, we began to consider what other social determinants of LLIN use we should investigate related to our project.

**Focus Group Discussions**
With this information in mind, we planned our FGDs, modified with HCD exercises, to be held in Ghana. Nine FGDs were conducted in March 2017, in the Ashanti, Greater Accra, and Western regions of Ghana. Our vision was to facilitate a series of activities designed to capture the thoughts, feelings, and opinions of urban and rural middle-class Ghanaians related to malaria prevention, LLINs, and their consumer preferences (e.g., preferred colors for the home and bedroom). People participated in a variety of exercises designed for this purpose. For example, in one exercise, participants were asked to rank order their preferences of various accessories that might be bundled with an LLIN (e.g., adhesive hooks), using a stack of cards with images of the accessories on them. In another experiential exercise, stations were created to observe people interacting with actual LLINs of different sizes and designs that had been hung around the venue. We noted common behaviors such as cautious sniffing of the LLINs and continuous patting and “fluffing” of the LLIN from the inside. The information gathered from these activities coupled with data collected from our FGDs helped us better understand how we might improve upon existing LLINs by creating designs that would be appropriate, as well as desirable, for this group of people.

Although the FGDs were designed and facilitated as a part of a cohesive suite of HCD activities executed throughout the overarching project, we have chosen to highlight these specifically because they yielded some important insights that subsequently influenced the features of our preliminary LLIN design prototypes. HCD is inherently transdisciplinary, and in our research and practice, we have found that its execution can be highly dependent on the implementer. From our perspective, the benefit of using a method like FGDs in an HCD context is that it can be used more flexibly than approaches that are wedded to a single disciplinary approach or formal research guidelines that place an overarching priority on standardization and protocol. For us, this meant allowing facilitators and interviewers the freedom to explore and follow unexpected or interesting threads in the conversations. It also allowed us to incorporate less conventional activities (e.g., expressing oneself through creative activities).

Importantly, the FGDs yielded insights that influenced the design of our LLIN prototypes.

Through a stakeholder analysis, we visualized potential partners and opportunities as well as existing barriers.
Recruitment of Participants
We used a nonprobability sampling approach, or convenience sampling, to recruit participants for the FGDs. The total sample included 78 participants (51 adults and 27 students).

The study included 2 main target groups, namely:
- Middle-class adults (ages 18–59) who possessed the disposable income necessary to afford purchasing LLINs from the retail market.
- Senior high school students (ages 15–17) in boarding schools who were required to bring LLINs to their boarding houses, as indicated on their prospectus.

A standard screener, in the form of a questionnaire, was administered to identify participants who met the qualifications for participation, which involved previous experience using bed nets, income level (lower-middle to upper-middle class), profession, employment status (e.g., employed, not employed, boarding school student), and age. This screener helped us achieve adequate representation of our groups of interest. Participant demographics are presented in Table 2.

Procedures
Our FGDs employed a conversational style of group interviewing to explore issues related to experience with malaria, perceptions of malaria, methods for malaria prevention, and facilitators and inhibitors of LLIN use. Each FGD had 2 facilitators, one of whom was proficient in the local language, and 2 notetakers. Facilitators were provided with loose scripts to ensure that certain topics were covered in all sessions. FGDs were conducted in English unless it was decided that the local language was more appropriate.

Our FGDs included the following exercises. Participants were observed and their reactions and comments were captured via audio recorder and by a notetaker during discussions and on video and by a notetaker during experiential activities (with participant permission per Institutional Review Board requirements):
- **Discussions**: Participants were asked to share their thoughts, feelings about malaria, net use, and current net designs. They were also asked for their suggestions on what design aspects would improve their chances of buying and using an LLIN.
- **Color/design preference activity**: Participants indicated their preferred color and design preferences for a bed net. They were provided with a black and white drawing of a bed net (various shapes) and asked to use art supplies to indicate any preferred color choices or designs (e.g., logos, patterns).
- **Net material preference activity**: Participants were given swatches of 3 different materials (polyester, polyethylene, or polypropylene) currently being used for LLINs and asked to describe their reactions to touching them and discuss their preferences and perceptions of the materials.
- **Bundling preferences**: Participants were shown cards with a selection of hypothetical accessories that could be bundled with a bed net

FGDs enabled exploring experience and perceptions of malaria, methods for malaria prevention, and facilitators and inhibitors of LLIN use.

<table>
<thead>
<tr>
<th>TABLE 2. Demographic Summary of Focus Group Discussion Participants, by Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group</strong></td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
</tbody>
</table>
to gauge their desire and preferences for such products. They included items such as adhesive hooks and small fans. These items were chosen because they emerged in our research and initial interviews as potentially attractive products to bundle with an LLIN that might encourage purchase. Participants were asked to sort the cards in order of preference and discuss their reasoning aloud. Card sorting was used because it can help participants prioritize preferences more easily than using pen and paper or a drop-down menu on an online survey.

- **LLIN experiential activity:** Nets of various shapes and sizes were set up around the workspace. Participants were observed while they examined, entered, spent some time inside, and exited the nets. They were invited to describe any of their feelings about the experience to facilitators and notetakers.

- **Time to reflect and conclusions:** Participants were given an opportunity to ask any remaining questions or share any final thoughts before departing.

### Data Analysis

Data captured from the FGDs were transcribed and analyzed together with additional field notes taken by the facilitators and notetakers. We used the micro-interlocutor analysis framework for qualitative analysis of these data to capture emergent themes. Compared with other methods of FGD data analysis, this framework allowed us to consider the voices, or lack thereof, of all participants in the insight generation process by analyzing the answers given and not given by each participant instead of seeking or focusing primarily on points of group consensus.

### RESULTS

Three main constructs—convenience, comfort, and aesthetic—emerged from our analysis regarding key barriers to LLIN use as described by our participants (Table 3). The barriers were related to use of a double-size rectangular LLIN with 4 hanging points, which is the most common type of LLIN distributed in Ghana through the public sector. With regard to convenience, several participants described the process of hanging as well as entering and exiting the LLIN as challenging, stressful, and/or tedious. Another issue was the lack of access to things around the bedroom, such as eyeglasses, phones, bibles, and condoms, while using the LLIN. Participants were also concerned about discomfort associated with using the LLIN. Specifically, use of LLINs was considered to make people feel even hotter in an already warm climate as well as to leave users feeling confined within a small space. Multiple participants mentioned that using an LLIN was not conducive to having sex for similar reasons. People also complained about the polyethylene material of the LLIN saying it was “hard” and rough to the touch, and some expressed concerns regarding the insecticide treatment and perceived side effects associated with its use (e.g., burning sensations in the eyes or on the skin). Finally, participants also discussed how LLINs were aesthetically incompatible with their desired bedroom and home décor. Many discussed their thoughts on how to improve the look of LLINs including suggestions for additional colors, shapes, and hanging mechanisms, among other design modifications.

The participants provided us with a great deal of information about their lifestyles and LLIN preferences. More importantly, their responses support our view that an opportunity exists to modify the standard LLIN design to satisfy the contemporary consumer preferences of the Ghanaian middle class. We used these emergent constructs as the basis for developing differentiated LLIN designs that would be more desirable to this target population.

### Key Insights

We found that LLIN usage gaps among our middle-class Ghanaian participants may exist to some degree because of evolving consumer preferences. From what we gathered, their current preferences reflect a desire for things that match contemporary sensibilities (e.g., convenience, comfort, and personal style) and reflect an upwardly mobile lifestyle. We found that trade-offs are made on functional effectiveness and affordability for a mosquito control product that is easier to use. This finding may explain why an overwhelming majority of our target population preferred non-LLIN malaria control products (e.g., aerosol insecticide sprays and coils) despite acknowledging that LLINs are more effective for malaria prevention and less expensive. As previously mentioned, these alternative forms of mosquito control are quite prevalent among this demographic and compete directly with LLINs because they are affordable, convenient, and, we posit, perceived to be a more modern approach to malaria prevention than LLINs, which carry a different connotation. We argue that in order for
TABLE 3. Key Barriers to LLIN Use According to Focus Group Discussions

<table>
<thead>
<tr>
<th>Construct</th>
<th>Barriers to Use</th>
<th>Selected Quotations</th>
<th>Examples of Suggested Improvements</th>
</tr>
</thead>
</table>
| Convenience        | Hanging:                                                                      | “Bed net installation is stressful.”  
                       | • Tedium to affix or hang because of multiple (at least 4) hanging points         | — Male Adult, Rural, Ashanti Region                 |
|                    | • Challenges with finding accessories needed for hanging (e.g., nails, hooks, poles, and strings) |                                                                                     | LLIN with minimal hanging points (e.g., single point hang) |
|                    | • Concerns regarding defacement of bedroom walls by drilling nails or hooks that serve as hanging points |                                                                                     | Provision of hooks (regular or adhesive) and string with LLIN |
| Entry and exit:    | • Not easy to enter and exit the LLIN—users have to lift them over their heads each time they enter or exit as well as tuck them under the mattress once inside | “In the middle of the night, when I have to pee, I have to think hard about it, should I wait until morning, or should I go through the stress of getting out of my net?”  
                       | — Female Student, Urban, Greater Accra Region                                       | — Female Adult, Rural, Greater Accra Region         |
|                    | Access to personal items:                                                     | “As a mother with a baby, I can keep baby pampers in the pocket.”                    | A pocket inside the LLIN for storage of items that may be needed during the night |
|                    | • When inside the LLIN, users are restricted from accessing nearby personal items that are outside the net (e.g., things on a bedside table, such as bible, phone, eyeglasses, water, and personal items) | — Female Adult, Rural, Greater Accra Region                                                 | Improved entry and exit also addresses this barrier |
| Comfort            | Restrictive space:                                                            | “I am married and I can’t see myself playing games with my wife in a bed net. Even when I was single the bed net was not spacious for me alone. Imagine now that I am married, it won’t work unless it is made spacious, so we can feel free in it.”  
                       | — Male Adult, Urban, Ashanti Region                                                 | — Male Adult, Rural, Western Region                |
|                    | • The feeling of being closed-in (e.g., claustrophobic sensation) or lack of space within the LLIN |                                                                                     | Rectangular LLINs as they are perceived as more spacious |
|                    | • Particularly an issue for sexually active couples                            |                                                                                     | LLINs in various bed sizes                          |
|                    | Heat:                                                                         | “The weather, the weather, the weather. It is just hot for bed nets! Especially when you use those hard [polyethylene] nets.”  
                       | • Weather conditions are too hot to use an LLIN comfortably                       | — Male Adult, Rural, Western Region                |
|                    | • Perception that LLINs increase the heat felt while sleeping                  |                                                                                     | Provide a small fan to reduce the heat (e.g., Briët et al.) |
|                    | Material texture:                                                             | “The [polyethylene net] is hard and can give skin rashes.”                          | Use polyester material for LLINs as it is perceived to be softer and less hot |
|                    | • The material texture of polyethylene LLINs are perceived as both rough and hot | — Male Student in Urban, Ashanti Region                                               |                                                     |
|                    | Insecticide treatment:                                                        | “I didn’t like the [polyethylene net] because it feels like rubber, so if you use it, there would be a lot of heat.”  
                       | • Adverse reactions to the insecticide in LLIN (e.g., skin irritation and itchiness)| — Female Adult, Urban, Greater Accra               | Improved messaging about proper care of LLINs before first use (e.g., air out LLIN for at least 24 hours before using) |
|                    | Aesthetics                                                                    | “The real problem is the chemicals in the net. If we use it and it makes us uncomfortable then we stop using it.”  
                       | • LLINs detract from bedroom décor by creating a cluttered look in the bedroom    | — Female Adult, Urban, Greater Accra               |
|                    | • Limited options of LLIN styles and sizes that complement bedroom décor       | “The nets make your room look ugly. We are forced to use it because it protects us against mosquito bites. We need fancy nets, with a variety of colors we can choose from, and ones that are nice to hang so we can do away with all the poles and nails in the walls.”  
                       | — Female Adult, Rural, Greater Accra                                              | — Female Adult, Rural, Greater Accra               |
|                    |                                                                                |                                                                                     | An LLIN design that is more sleek (e.g., minimal hanging points) |
|                    |                                                                                |                                                                                     | LLINs available in a variety of sizes and colors    |

Abbreviation: LLIN, long-lasting insecticidal net.
LLIN use to increase among this population, the design of LLINs needs to change. The new design(s) must be convenient, comfortable, and aesthetically pleasing, while also maintaining the key functionality of protecting against mosquito bites.

Changing or tweaking certain LLIN features could aid in demedicalizing their use, thus positioning LLINs to become something more desirable to use in the home. To date, LLINs have been positioned as a tool almost entirely for malaria prevention purposes and have been distributed with a 100% subsidy (free of charge) like many other health commodities. Our goal in introducing the concept of differentiated LLINs was to create a product that is convenient for use, similar to the non-LLIN mosquito control products, and also to position it as something that would not detract from, and might even improve, one’s interior design. Additionally, LLINs currently on the market are primarily sold through pharmacies with little demand, and our research suggests that there might be a repositioning opportunity by selling LLINs in other, nonmedical, retail outlets (e.g., supermarkets, convenience shops).

Prototype Development and Testing

With an understanding of these 3 barriers to LLIN use, the study team, along with members of the target market (a panel of middle-class Ghanaians), devised and tested different LLIN attributes that address these drivers. In keeping with HCD best practices, the insights yielded from our in-country work were thoroughly reviewed by our project team and became fodder for several brainstorming sessions on how to best meet the stated consumer preferences of our target populations within our constraints (e.g., maintaining vector control efficacy, economic factors). Our intention was to use our learning to create prototypes that addressed the barriers and pain points indicated by our participants. After much discussion, we prioritized what we felt were the most important (and feasible) LLIN design features, which ultimately included a more convenient way to hang the net, a more attractive silhouette, and a zipper that allows the user to enter and exit with ease while still maintaining a sealed, mosquito-free space. A separate discrete choice experiment confirmed the attractiveness of these attributes by capturing the target population’s willingness to pay for LLINs with said preference-congruent attributes.17

We have since shared our consumer insights and preliminary ideas for new design features with current manufacturers globally who supply LLINs in Ghana. We hope these partners take this information into consideration as they make decisions about current and future LLIN supply, demand, and marketing and will pursue pilot testing of new net designs for the private-sector retail market in Ghana. Should that opportunity come to fruition, we intend to support efforts to test the ideas in-country to continue improving designs for the maximal benefit of our target population. This iterative approach to solution refinement directly reflects the tenets of HCD.

**DISCUSSION**

Human centeredness should not be thought of as a binary designation but rather a feature or quality upon which to pursue continuous improvement. In other words, achieving human centeredness in public health services and interventions does not have to have a ceiling. Moreover, just because HCD is suggested as an approach does not necessarily mean that the target of these efforts was not originally created with key stakeholders or users in mind. For example, in the case of this project, it was suggested that HCD could take a medically effective product and make it more desirable to use among a population that had experienced an immense amount of societal and economic change since the product was first designed. An HCD approach was pursued because it was appropriate and ideal for the type of information we needed to achieve this goal.

One of the most prominent lessons we learned from this work is the importance of contextualizing the use of public health interventions so they are tailored to specific settings and populations. It is unlikely that any single LLIN design can serve all people equally well. Our work confirmed what we already knew—even within a single country like Ghana, people are not monolithic. To serve the needs of the middle class in the urban and rural areas in the course of catalyzing a commercial market for LLINs, interventions should be considered in the overall context. Specifically, the LLIN can be thought of not only in terms of vector control and malaria prevention but also as a consumer product that may be used or not used because of how it makes people feel about themselves and their homes. This approach is important for public health professionals because the LLIN is currently the most effective form of malaria control, but it has competition from other, less medically effective products that are viewed as consumer/home goods.

We also discovered that prevention of malaria competes with other values and objectives, such as
People care about their health, but they also care about sex, romance, religion, their homes, comfort, and other things.

Convenience, comfort, and aesthetics. When researchers consider LLINs, they may not necessarily view them in relationship to use of bibles and condoms kept on nightstands; however, our target population does. People care about their health, but they also care about sex and romance, their religious practice, the look and feel of their homes, and personal comfort, among other things. We saw these priorities as potentially being in competition, so it is important to consider how to accommodate them in new LLIN designs. We believe these kinds of discoveries will more readily emerge from taking an HCD approach because a more traditional research approach may require a narrower focus or greater adherence to standardization and clear-cut methods. HCD helped us to identify the needs and wants of a population that has experienced a great deal of change in the last couple of decades. It also helped us to glean consumer preferences for a product about which there is almost minimal knowledge of consumer preferences and to design a medical product that may also be commercially desirable. Human centeredness was at the core of each of these questions. The flexible, personal, user-centered approaches of HCD made sense for us and provided advantages over other traditional research methods.

The introduction and execution of self-expressive and experiential exercises within the context of FGDs significantly improved the richness of our data. The inclusion of these activities helped set a tone for the day of work together that promoted candid sharing and creativity. Additionally, the underlying philosophy of HCD as an experimental, iterative form of problem solving encouraged the project team and the stakeholders to play with their ideas and collaborate across groups and teams. Gleaning the same insights or soliciting the same level of engagement through traditional approaches is difficult to imagine.

We were surprised at how pleased our participants were to be consulted on the design of a product for them. One female participant in the rural area of Ashanti Region remarked:

*We always have free bed nets given to us, or have to buy [sic] whatever bed nets shops have available to sell, yet no one has ever taken the time to come ask us what we want in a bed net. So thank you for asking us that today and taking our opinions into consideration!*

For us, this comment was noteworthy. We initially decided upon an HCD approach to ensure fidelity to our target population’s needs and interests. We did not necessarily expect that engagement in the process would also be meaningful to participants.

For the portion of our work that we chose to highlight in this article, our project team selected an approach that aligns closely with qualitative research methods. It made sense for us given our project goals and our team’s training and experience. It is important to note that HCD does not promote any single academic perspective, methodology, or approach over others. HCD is practiced differently depending on the composition and collective expertise of the team and the problem or challenge that it seeks to address. We believe this openness and flexibility can be a tremendous advantage for those who seek permission to explore innovative and/or cross-collaborative methods and approaches to their work. We believe our team’s experience can be learned from, but we do not submit our work as a template for all global health HCD projects.

The increasing number of public health projects that use HCD appears to connect more broadly to a larger trend of human-centered initiatives in health care. The concept of health and well-being today is more nuanced such that hospitals, care providers, medical technology companies, and others are realizing that they can no longer afford to focus solely on the primary health goal. They must also consider how social determinants of health influence behaviors and intervention uptake. For this project, we realized we had to expand our thinking beyond the preventative medical aspects of the LLIN in order to meet the research objective of catalyzing a commercial market for LLINs and ultimately the goal of increasing malaria prevention among our target population. For a health technology such as the LLIN to produce a benefit, it has to be used. Therefore, public health professionals should pay serious attention to the factors that drive or constrain use, including the perceptions and preferences of end-users.

**Acknowledgments:** We would like to recognize and thank all the individuals who participated in our human-centered design work, including the Private Sector Malaria Prevention project teams and affiliates in Baltimore and Accra, focus group discussion and workshop participants, and various public health and business partners who were consulted for their knowledge and expertise. We also thank Felix Nyamar-Fosu, Tetteh Ogum, Kathryn Bertram, and the URIKA research team for their contributions to this project.

**Funding:** This article was made possible by the generous support of the United Kingdom Department for International Development (DFID). The authors are responsible for the content. The content does not necessarily reflect the views of DFID or the government of the United Kingdom.

**Competing Interests:** None declared.
Using Human-Centered Design to Determine Preferences for LLINs

REFERENCES


Peer Reviewed

Received: August 2, 2018; Accepted: April 2, 2019


© Kim 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-18-00284
Cell Phone Counseling Improves Retention of Mothers With HIV Infection in Care and Infant HIV Testing in Kisumu, Kenya: A Randomized Controlled Study

Avina Sarna, Lopamudra Ray Saraswati, Jerry Okal, James Matheka, Danmark Owuor, Roopal J. Singh, Nancy Reynolds, Sam Kalibala

Tailored, one-on-one counseling delivered via cell phone was very effective in retaining mothers with HIV in care and in promoting infant HIV testing and antenatal and postnatal care attendance. The highest risk of loss to follow-up among women with HIV accessing PMTCT services was prior to delivery and then after infant HIV testing at 6 weeks. Challenges include continued limited access to cell phones, difficulty with reaching participants on the phone, and poor adherence to antiretroviral therapy for a substantial percentage of the population.

ABSTRACT

Background: We evaluated the effectiveness of a cell phone counseling intervention to promote retention in care and HIV testing of infants among women with HIV accessing prevention of mother-to-child services in Kisumu, Kenya.

Methods: Between May 2013 and September 2015, we recruited 404 pregnant women with HIV who were between 14 and 36 weeks of gestation and randomly assigned them to the intervention (n=207) or control arm (n=197). Retention was assessed at delivery and at 6 and 14 weeks postpartum. We also measured uptake of infant HIV testing. The intervention comprised a fixed protocol of counselor-delivered phone calls to provide one-to-one need-based support. The number of calls made varied depending on when participants presented for antenatal care services; the maximum number was 42. The control group received routine care. We evaluated retention at 3 time points using the complementary log-log regression model taking into account factors associated with retention and loss to follow-up time. We calculated the incidence rate for HIV transmission among infants and used binary logistic regression to identify predictors of HIV infection among infants.

Results: Participants attended on average 63% of the required number of counseling calls during the study period. Retention was higher in the intervention arm than the control arm at delivery (95.2% vs. 77.7%, respectively); 6 weeks postpartum (93.9% vs. 72.9%, respectively); and 14 weeks postpartum (83.3% vs. 66.5%, respectively) (P<.001). The counseling intervention (hazard ratio [HR]=0.29; 95% confidence interval [CI]=0.12, 0.69) and positive health perceptions (HR=0.99; 95% CI=0.98, 1.00) were associated with lower hazards of being lost to follow-up. HIV testing of infants was higher in the intervention than control arm (93% vs. 68%, respectively; P<.001). In total, 9 of 308 (2.9%) infants tested positive for HIV infection (incidence rate=0.39 infections/100 infant-weeks). Medication Possession Ratio (MPR) >90%, used to assess adherence to ART, was associated with lower odds of a positive HIV test among infants (adjusted odds ratio=0.20; 95% CI=0.04, 0.99). Attendance at antenatal and postnatal care visits was higher among participants in the intervention arm than the control arm.

Conclusions: The one-on-one tailored counseling delivered via cell phone was effective in retaining mothers with HIV infection in care and promoting uptake of infant HIV testing and antenatal and postnatal care services. Phone counseling offers a practical approach to reach and retain pregnant women with HIV infection and postpartum mothers in care, but greater emphasis on collection of medications and adherence is required.

INTRODUCTION

The Global Plan Towards the Elimination of New HIV Infections Among Children by 2015 and Keeping Their Mothers Alive (Global Plan) was launched in 2011. The 2015 Global Plan progress report for Kenya revealed that 70% of pregnant women living with HIV received antiretroviral therapy (ART).1 There were 13,000 new infections among children in 2014, with a mother-to-
We evaluated the effectiveness of structured cell phone counseling on maternal retention in care and uptake of infant HIV testing in Kisumu, Kenya.

Child transmission rate of 5% at 8 weeks and 17% when including breastfeeding. In addition, the 2013 Kenya AIDS Indicator Survey reports an HIV prevalence of 6.1% among women who had a birth in the 5 years preceding the survey. Mother-to-child transmission remains a challenge, with poor retention rates for mothers in care and consequently less-than-complete uptake of testing for early infant diagnosis (EID) between 6 and 8 weeks of age.3

A series of recent systematic reviews4–6 report a variety of interventions implemented in low-middle-income countries to promote retention of mother-baby pairs and uptake of HIV testing and ART among exposed infants. Vrazo et al. (2018)6 found interventions focused on antenatal care (ANC) and ART integration, family-centered approaches, and the use of lay health care providers to be demonstrably effective in increasing service uptake and retention in care. Geldsetzer et al. (2016)5 report that overall the evidence base for interventions to improve postpartum retention in HIV care is weak, except for some evidence that phone-based interventions have a positive impact during the first 3 months postpartum.5 Evaluations of interventions such as male involvement,7 mother-to-mother peer groups/mentor mother programs,8 community health worker support,9 and continuous quality improvement10 have provided varying results. While the mentor mother intervention was successful in improving mothers’ retention in care (retention was 61.9% in the mentor mother arm compared with 24.9% in the control arm), the male involvement, community health worker support, and continuous quality improvement interventions did not show any effect. Other quality improvement projects using rapid results initiative11 and health system redesign,12 as well as other community health worker interventions13,14 using individualized community-based follow-up, were successful in improving ART initiation among mothers with HIV and their infants, but did not specifically examine retention rates of mothers in care, especially during the antenatal period and beyond 6 weeks postpartum.

With the widespread availability of mobile phones there is increased interest in the use of technology-based methods to improve health services uptake for better health outcomes among mothers with HIV and their exposed children. Ambia et al.’s systematic review of 34 studies12 includes 5 studies (2 of which are randomized) that evaluated mobile phone-based interventions that showed a significant increase in the uptake of EID at 6 weeks.3,15–18 Three studies conducted in Kenya have evaluated mobile phone-based interventions to improve retention, adherence to treatment, and uptake of HIV testing for infants; however, almost all of these interventions have used short message service (SMS) as reminders3,15,16 that show an improvement in the uptake of infant HIV testing but not in the retention of mothers in HIV care. Counseling support delivered via cell phones has been shown to improve adherence among people living with HIV who are on ART.19,20 One-on-one counseling interventions based on self-regulation have been evaluated for adherence to ART among people living with HIV in the United States.19,20 Similar counseling interventions have also been used for smoking cessation among people living with HIV in the United States.21 There are, however, no studies that have used mobile phones to deliver one-on-one counseling support for pregnant women living with HIV and who are accessing prevention of mother-to-child transmission (PMTCT) services.

We evaluated the effectiveness of a structured cell phone counseling intervention, informed by behavioral theory and delivered by trained counselors, on maternal retention in care until 14 weeks after birth and uptake of EID/HIV testing in Kisumu, Kenya. The project was called the Healthy Mother Healthy Baby Project.

METHODS

Study Design and Study Sites

We conducted a parallel-group, unblinded, randomized controlled study among pregnant women living with HIV who were accessing PMTCT services in Kisumu, Kenya. The primary objective of our study was to determine whether a structured, counselor-delivered, tailored cell phone counseling intervention would increase (1) retention in care until 14 weeks postpartum, and (2) uptake of EID or infant HIV polymerase chain reaction (PCR) testing. As secondary outcomes, we examined HIV transmission among HIV-exposed infants and maternal attendance at ANC and postnatal care (PNC) services.

The study was conducted at 14 HIV treatment clinics providing PMTCT services in Kisumu County. High-volume sites were selected in consultation with County AIDS Control officials from a list of clinics providing PMTCT services under the AIDS, Population and Health-Integrated Assistance (APHIA) Plus Program in Kisumu, which is supported by the United States
Agency for International Development. All clinics provided similar PMTCT services per the national protocol with regard to provision of antiretroviral (ARV) medications and client follow-up. Between May 2013 and September 2015, pregnant women living with HIV were recruited and randomly assigned to the intervention and control arms using computer-generated random numbers. Participants were followed up to 14 weeks postpartum.

Eligibility Criteria
We invited pregnant women living with HIV who were between 14 and 36 weeks of gestation, aged ≥16 years, residing in Kisumu and planning to stay there for the next 12 months, willing and able to provide consent, and who had access to a cell phone (owned or shared) to participate in the study. Participants could be ART naïve or experienced (they were currently on ART or had received nevirapine for a previous pregnancy). Clinic nurses informed potential participants about the study, obtained verbal consent, and then introduced them to the study staff for completing consent and recruitment procedures.

Study Visits
All study visits were linked to routine maternal and child health services: monthly ANC visits before delivery (the national program recommends a minimum of 4 scheduled comprehensive ANC visits during pregnancy), PNC visit at 6 weeks after delivery (the national program recommends a minimum of 3 visits: the first, between 24 and 48 hours of delivery, the second between 7 and 14 days after delivery, and the third at 6 weeks after delivery), and infant immunization visits at 6, 10, and 14 weeks of age at the clinic (the Expanded Program on Immunization recommends bacille Calmette-Guérin (BCG)/polio/hepatitis B at birth and diphtheria, pertussis, and tetanus (DPT)/polio/hepatitis B/pneumonia at 6, 10, and 14 weeks). Participants completed a baseline interview upon recruitment and an endline interview at 14 weeks postpartum when they visited the center for completing the last of the primary vaccinations for infants. Monthly data were collected when clients visited the clinic for collecting monthly ART medications. Standardized data collection tools, staff training, and regular supervision ensured that study activities were uniform across sites.

Description of the Intervention

Standard Care
All newly diagnosed pregnant women living with HIV (ART naïve) and those who became pregnant while on antiretroviral therapy (ART experienced) received routine HIV counseling from ART clinic-based counselors. The counseling included information on the risk of HIV transmission to the infant, the role of ART in PMTCT, the importance of adherence to treatment, disclosure and partner testing, institutional delivery, and infant HIV testing at 6 weeks postpartum. All participants received standard ANC services, which included blood pressure and weight measurements, hemoglobin, syphilis and urine testing, tetanus toxoid immunization, and iron and folic acid supplements. All participants also received standard PNC services, which included mother’s check-up, HIV PCR testing for infants at 6 weeks postpartum, and routine immunization services. All centers had peer community health workers associated with the clinic to support clients and trace those who defaulted or missed visits.

Cell Phone Counseling Intervention
In addition to standard care, participants in the intervention arm received one-on-one individualized counseling, delivered via cell phone by 5 trained counselors based at a central study office. The counseling was drawn from the Self-Regulation Theory, which is a system of conscious personal management that involves the process of guiding one’s thoughts, behaviors, and feelings to empower patients to recognize their problems and find solutions.21 The sessions were structured to consist of 2 phone calls during the first week of starting PMTCT services, followed by 1 call/week until the participant delivered (maximum of 26 calls), followed by 2 calls during the first week after delivery and 1 call/week for 14 weeks thereafter (maximum of 16 calls) (Figure 1). The number of calls during the antenatal period varied between participants depending on when they presented for ANC services (between 14 and 36 weeks of gestation).

Training of Counselors: The counselors were trained HIV counselors who had at least 3 years of experience in HIV counseling at various testing centers and had completed at least high school education. The counselors received a 10-day training on the intervention that included theoretical background of the intervention, training on counseling techniques, role play with colleagues and trainers, and practice sessions with
volunteers with HIV infection. During the first month, counselors were required to debrief with the program coordinator after each call and receive feedback. Thereafter, the counselors continued with a weekly group discussion on problem cases.

**Intervention Sessions:** After completing recruitment procedures, the research staff put the participants in touch with a study counselor via phone; the participants never met their counselor face-to-face during the entire study period. Participants were required to use their own phones, including a phone shared with a family member or friend. Counselors and patients decided mutually convenient times for the calls.

The first session focused on illness representation and problem identification. This was followed by the development and execution of a response plan and evaluation of coping strategies over follow-up sessions. Perceptions or contextual situations that could pose an impediment to ARV adherence or retention in care were identified and participants were encouraged to think about their experiences, interaction with others, sources of information, and cognitive and emotional processes that contributed to their perceptions. Participants were then encouraged to discuss strategies on how to manage their perceptions. Through this process, the counselors introduced replacement perceptions and alternate behaviors. The counselors helped participants address their areas of concern by providing targeted action plans, setting realistic goals, and assessing progress during the next follow-up call.

During the early antenatal period (14–32 weeks), counselors focused on the importance of adherence to treatment for their own health and to ensure their baby was born HIV-free. Partner disclosure, partner testing, stigma issues within the family/community, distance to ART centers, and travel constraints were assessed and participants counseled. The focus shifted to emphasize retention in care, institutional delivery, and the need for initiating nevirapine for the baby in the late antenatal period (32–40 weeks) while continuing to emphasize adherence. During the postnatal period (0–15 weeks postpartum), the counselors discussed nevirapine for the infant, infant feeding (exclusive breastfeeding), PCR testing of the infant at 6 weeks, completion of the primary immunization schedule of vaccines, and family planning for the mother while continuing to emphasize the need to continue ART and remain adherent. The Box details the topics covered during the calls.

Participants could make additional need-based calls to the counselor during working hours on weekdays to address concerns or queries. The calls enabled participants to have frequent, personalized, one-on-one contact with a health care professional without visiting the health facility. Data were collected on the frequency and duration of calls made, number of attempts made to reach the client, and reasons for unsuccessful calls. All study participants received a baby gift pack containing soap, baby oil, and disposable napkins when they came for their PNC visit at 6 weeks.

**Data Collection and Study Variables**

Data were collected using structured questionnaires administered by research assistants in Swahili or Luo. Variables were categorized as follows: **education** as never attended school, received primary education, or attended secondary or university education; **marital status** as never married, married or cohabiting, or divorced/separated/widowed; **living arrangements** as lives alone, lives with partner/husband and children, or lives with others. **Pregnancy duration** at recruitment was categorized as 14–28 weeks or 29–36 weeks; **time since HIV-positive status** as...
1 year or less, 2–4 years, or 5 years or more; and partner/spouse’s HIV status as positive, negative, or unknown. Participants were considered ART naïve if they were diagnosed positive but had never received ART and experienced if they became pregnant while on ART. PMTCT treatment regimens were categorized as Option A or only AZT (zidovudine) for the mother, or Option B or combination of 3 ARVs for the mother; infants received nevirapine under Option A and nevirapine or AZT under Option B. Depression was assessed at baseline and endline using the Center for Epidemiologic Studies Depression (CES-D) scale, a 20-item validated scale. The scale has a possible range of scores of zero to 60 with higher scores indicating the presence of more symptomology. Depression was categorized as no depression if scores were <16 and depression if scores were ≥16. Perceived stigma was assessed at baseline and endline using a 16-item scale (Cronbach’s alpha of adapted scale: 0.81), derived from Berger’s HIV stigma scale25 that has been used in other studies in Kenya. The scale covered 4 domains: disclosure concerns, negative self-image, concerns about public attitudes, and personalized or experienced stigma. Total scores (range: 16–64) were categorized as low (16–40), moderate (41–52), or high (53–64) stigma.

**General health perception** was assessed using the Health-Related Quality of Life tool used by AIDS Clinical Trials Group studies. The tool examines perceptions about general health; resistance to illnesses and health outlook; physical, social, role, and cognitive functioning; and pain. Item scores in each scale are summed to compute raw scale scores that are then transformed to a 0 to 100 scale. Higher scores are indicative of better health functioning. Scores were categorized as above average (61–100) or average or below (≤60).

**Adherence** was assessed using the Medication Possession Ratio (MPR) derived from pharmacy refill information, collected from pharmacy registers, and recorded as a percentage. MPR=Number of days participants had supply of medications/Number of days in the study. For analysis, MPR was dichotomized as ≥90% or <90%.

**Retention in care** was assessed at 3 time points: at delivery, 6 weeks postpartum, and 14 weeks postpartum. Participants who delivered at the health facility where they received PMTCT services, or at another health facility, or for whom there was information of a home delivery and pregnancy outcome were considered retained.

### Content Focus of Counseling Calls

<table>
<thead>
<tr>
<th>Initial antenatal period (14–32 weeks gestation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Personalized problem identification (first session)</td>
</tr>
<tr>
<td>• Importance of ART for PMTCT and infant outcomes</td>
</tr>
<tr>
<td>• Adherence to treatment, including monthly collection of medications and timely and regular intake of medication</td>
</tr>
<tr>
<td>• Discuss the importance of retention in care, including visit attendance</td>
</tr>
<tr>
<td>• Discuss partner involvement, HIV status, disclosure, and testing</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Late antenatal period (32–40 weeks gestation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Emphasize retention in care</td>
</tr>
<tr>
<td>• Emphasize institutional delivery</td>
</tr>
<tr>
<td>• Introduce infant feeding, including exclusive breastfeeding</td>
</tr>
<tr>
<td>• Discuss nevirapine initiation for baby</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Postnatal period (0–14 weeks postpartum)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Confirm initiation and continuation of nevirapine for baby</td>
</tr>
<tr>
<td>• Infant feeding, including exclusive breastfeeding</td>
</tr>
<tr>
<td>• Infant HIV PCR testing at 6 weeks</td>
</tr>
<tr>
<td>• Completion of immunization, including BCG and polio at birth, DPT and polio at 6, 10, and 14 weeks</td>
</tr>
<tr>
<td>• Emphasize retention in care</td>
</tr>
<tr>
<td>• Introduce discussion on confirmatory HIV testing at 18 months</td>
</tr>
<tr>
<td>• Discuss family planning needs of the mother</td>
</tr>
</tbody>
</table>
at delivery. Participants who completed their 6-week PNC visit or had their baby tested for HIV (PCR test) or had the baby immunized at 6 weeks were considered retained at 6 weeks postpartum. Participants who had their baby immunized at 14 weeks were considered retained at 14 weeks postpartum. Participants with stillbirths and infant deaths prior to time points 6 weeks and 14 weeks postpartum were excluded from the analysis.

Data were collected on HIV testing of the infant. In Kenya, the national program requires infants born to mothers living with HIV to undergo HIV PCR testing at 6 weeks after birth. We collected information on HIV PCR testing undertaken any time between 6 and 14 weeks postpartum from the child health register. Data on attendance at ANC and PNC services, including infant immunization, were collected from the maternal and child health register. Monthly ANC visits coincided with the ARV pick-up from the pharmacy; the number of ANC visits during the study period varied depending on when the participant registered for ANC.

Counseling call details, such as number of calls made and duration of each session, were recorded by the counselors and verified from itemized monthly statements.

Statistical Analysis
Data were entered using the Census and Survey Processing (CSP) software program (U.S. Census Bureau and ICF Macro) and analyzed using Intercooled Stata 10.0 (Stata Corporation, College Station, TX). Patients were analyzed within the group to which they were originally assigned. Unpaired Student’s t test and Mann-Whitney U tests compared continuous variables with normal and non-normal distributions, respectively, and Pearson’s chi-square test was used to compare categorical variables. We compared sociodemographic characteristics, HIV testing and disclosure, pregnancy history, and ART use between intervention and control participants to assess the effectiveness of randomization.

We had 2 primary outcomes for this study: retention in care until 14 weeks postpartum and uptake of HIV PCR testing in the infant (EID). For the first outcome, we compared participant retention at 3 discrete time points in the study period: at delivery, 6 weeks postpartum, and 14 weeks postpartum (the endpoint coinciding with the end of the first set of primary immunization for infants). To evaluate retention at 3 time points while taking into account factors associated with retention and time to lost to follow-up (LTFU), we used a complementary log-log regression model, which is an alternative extension of the proportional hazard model for discrete time survival analysis. We calculated the complementary log-log of the hazard function at the 3 defined time points using the following model:

$$\log\left[-\log\{1 - \lambda(t_j|x_i)\}\right] = \alpha_j + \beta_j x_i$$

where:
- \(x_i\) is the vector consisting of sociodemographic, HIV–related, and study group variables for individual \(i\)
- \(\beta_j\) is the covariate matrix
- \(\lambda(t_j|x_i)\) is the hazard function for individual \(x_i\) at time point \(t_j\)
- \(\alpha_0=\log\left[-\log\{1 - \lambda(t_0|x_0)\}\right]\) is the complementary log-log of the baseline hazard

From the above model, we calculated the hazard ratio of being LTFU for an individual ‘i’ compared to the reference category at time point ‘j’ using following equation:

$$\frac{\lambda(t_j|x_i)}{\lambda(t_j|x_0)} = 1 - e^{-\alpha_0 + \beta_j x_i}$$

For each predictor variable, the baseline models were controlled for the time variable (the 3 time points), age, education, and marital status. We also examined the interaction effect of the predictor variable with the time variable. In the results, we display the interaction term only when found significant \((P<.05)\). The final model is a multivariate model controlling for age, education, and marital status. This model includes only the variables that were statistically significant in the baseline model.

Pearson’s chi-square test was used to compare the uptake of EID between the intervention and control arms. The incidence rate for HIV transmission among infants was calculated over the time period from birth to date of HIV PCR test by dividing the number of new infections by the total weeks of exposure. We used binary logistic regression to identify the predictors of HIV infection among infants.

We compared the uptake of ANC and PNC services among participants in the 2 groups using Pearson’s chi-square test. We also provide relative risk ratios (RRRs) for not attending at least 50% of the required visits (the number varied depending on when the participant registered for ANC), not taking the complete ANC package, not completing
3 PNC visits, not attending the 6-week PNC visit, and not delivering at a health facility in the intervention arm compared with the control arm.

**Ethical Considerations**

The study was approved by the Kenyatta National Hospital, University of Nairobi Ethics and Research Committee and the Institutional Review Board of the Population Council. All participants provided written informed consent.

## RESULTS

### Baseline Characteristics of Subjects

A total of 2,176 pregnant women living with HIV were screened at 14 PMTCT-ART centers in and around Kisumu County. Among those initially screened, 564 women refused to participate, 333 agreed to participate at a later date but never returned to complete recruitment procedures, and 875 were ineligible per the eligibility criteria, including 355 women who did not own or have access to a cell phone. In total, we recruited 404 pregnant women living with HIV; 207 were randomly assigned to the intervention arm and 197 to standard care (Figure 2).

There were no significant differences in sociodemographic and pregnancy-related characteristics between the intervention and control groups at baseline (Table 1), indicating effective randomization. In both groups combined, 57% of women were ART naïve, 21% had received ARVs for a previous pregnancy, and 22% were on ART for treatment of their infection (CD4 cell count <500 cells/ml). Two-thirds of the women had been diagnosed with HIV infection in the past year and 36% had a spouse with HIV while nearly half (47%) did not know the HIV status of their spouse/partner.

### Exposure to Intervention

Cell phones were not provided to the study participants; 68% of the intervention group participants used their own cell phone, 19% reported using their spouse’s cell phone, 8% used another family member’s phone, and 5% relied on friends. Reaching participants via phone calls was challenging and counselors had to make multiple calls to reach clients. It took an average of 4.8 call attempts (standard deviation [SD]=7.9) to make a successful call to participants. Overall, participants attended an average of 63% (SD=24.6) of the required number of counseling sessions via phone calls during the study period, with a counseling session lasting an average of 9.2 minutes (SD=7.9). Just over one-third (37%) of participants attended more than 75% of the required number of sessions/calls (Table 2). The average duration of the calls was higher among participants who attended fewer calls. Further, the average duration of counseling sessions was longer for participants with symptoms suggestive of depression (CES-D scale score≥16) compared with those without depression (10.8 minutes vs. 8.4 minutes, respectively; \( P < .001 \); data not shown).

### Effects of Cell Phone Counseling on Participant Retention

Participant retention in care was significantly higher in the intervention arm than the control arm at all 3 time points— at delivery: 95.2% vs. 77.7%; at 6 weeks: 93.9% vs. 72.9%; and at 14 weeks: 83.3% vs. 66.5% (Figure 3). All differences were significant at the \( P < .001 \) level. The highest dropout (44 participants; 22%) was observed before delivery among participants receiving standard care (Figure 2). This was followed by about 10.4% (20 participants) in the intervention arm who dropped out between 6 and 14 weeks postpartum (2 women with infant deaths and 1 woman who died after 6 weeks postpartum were excluded). In both arms, the dropout rate was lowest during the period between delivery and 6 weeks postpartum.

In the intervention arm, we observed a trend toward a linear relationship between retention and attendance at counseling calls, with higher retention among participants who attended a higher proportion of counseling calls (Table 2). The highest retention (92.7%) was observed among participants who attended between 51% and 75% of the calls. Among participants who were LTFU, the majority (70%) owned the cell phones they used. The use of shared phones did not have any effect on LTFU rates.

### Predictors of Loss to Follow-Up

In the baseline complementary log-log models, being employed (hazards ratio [HR]=1.69; 95% confidence interval [CI]=1.07, 2.33) and higher depression scores (HR=1.02; 95% CI=1.00, 1.03) were significant predictors of higher hazards of being LTFU (Table 3). On the other hand, knowing one’s HIV status for 5 or more years (HR=0.62; 95% CI=0.39, 0.97), having disclosed their HIV status to their partner (HR=0.68; 95% CI=0.44, 1.00), knowing that their partner has HIV compared with not knowing their partner’s HIV status.
(HR=0.62; 95% CI=0.38, 0.96), having better general health perception (HR=0.98; 95% CI=0.97, 0.99), and being in the intervention group (HR=0.22; 95% CI=0.11, 0.46) were significantly associated with lower hazards of being LTFU. Compared with the antenatal period, the lowest risk of being LTFU was in the 6-week postpartum period (HR=0.21; 95% CI=0.10, 0.43), followed by the period between 6 and 14 weeks postpartum (HR=0.76; 95% CI=0.49, 1.13; not statistically significant). Overall, participants in the intervention group had a lower hazard of being LTFU (HR=0.22; 95% CI=0.11, 0.46) and LTFU was lower during the 2 postpartum periods (delivery to 6 weeks: HR=0.21; 6 to 14 weeks: HR=0.76). However, the interaction term with time was significant for the study group; thus, between the 2 postpartum time periods, participants in the intervention group had a higher hazard of dropping out in the period between 6 and 14 weeks postpartum.

The final regression model shows that the conditional hazards of being LTFU during the postpartum period were significantly lower compared with the antenatal period: for 6 weeks postpartum (HR=0.27; 95% CI=0.10, 0.70) and for the 6–14-week period (HR=0.47; 95% CI=0.22, 0.89). Many of the variables that were significantly associated with LTFU in the baseline model lost their

**FIGURE 2.** Flow Diagram of Participant Recruitment and Follow-Up in the Healthy Mother Healthy Baby Project in Kisumu, Kenya (2013–2016)

- Pregnant women with HIV screened/reached (n=2,176)
- Participants consented to participate (n=404)
- Allocated to cell phone Counseling Intervention Arm n=207
  - Refusals: 2
  - Left Kisumu: 5
  - Died: 0
  - Untraced: 3
  - Participants followed up until delivery n=197
    - Still births: 9
    - Infant deaths: 3
    - Refusals: 0
    - Left Kisumu: 2
    - Died: 0
    - Untraced: 0
    - Participants completed 6 weeks postpartum (either EID or immunization) n=183
      - Infant deaths: 2
      - Left Kisumu: 14
      - Died: 1
      - Untraced: 3
      - Participants completed child immunization at 14 weeks postpartum n=160
- Allocated to Standard Care n=197
  - Refusals: 3
  - Left Kisumu: 14
  - Died: 0
  - Untraced: 27
  - Participants followed up until delivery n=153
    - Still births: 8
    - Infant deaths: 1
    - Refusals: 0
    - Left Kisumu: 4
    - Died: 0
    - Untraced: 3
    - Participants completed 6 weeks postpartum (either EID or immunization) n=137
      - Infant deaths: 0
      - Left Kisumu: 6
      - Died: 0
      - Untraced: 27
      - Participants completed child immunization at 14 weeks postpartum n=125

Abbreviation: EID, early infant diagnosis.
### TABLE 1. Baseline Characteristics of Participants Recruited in Kisumu, Kenya (2014)

<table>
<thead>
<tr>
<th></th>
<th>Intervention (n=207)</th>
<th>Control (n=197)</th>
<th>Total (N=404)</th>
<th>P Value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age, years, median (IQR)</strong></td>
<td>24 (22, 28)</td>
<td>25 (22, 29)</td>
<td>25 (22, 29)</td>
<td></td>
</tr>
<tr>
<td><strong>Educational level, n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No education or adult literacy</td>
<td>4/207 (1.9)</td>
<td>3/197 (1.5)</td>
<td>7/404 (1.7)</td>
<td>.72</td>
</tr>
<tr>
<td>Primary/preschool</td>
<td>133/207 (64.3)</td>
<td>134/197 (68.0)</td>
<td>267/404 (66.1)</td>
<td></td>
</tr>
<tr>
<td>Secondary or higher</td>
<td>70/207 (33.8)</td>
<td>60/197 (30.5)</td>
<td>130/404 (32.2)</td>
<td></td>
</tr>
<tr>
<td><strong>Current marital status, n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married/cohabiting</td>
<td>169/207 (81.6)</td>
<td>164/197 (83.3)</td>
<td>333/404 (82.4)</td>
<td>.31</td>
</tr>
<tr>
<td>Never married</td>
<td>25/207 (12.1)</td>
<td>16/197 (8.1)</td>
<td>41/404 (10.2)</td>
<td></td>
</tr>
<tr>
<td>Divorced/separated/widowed</td>
<td>13/207 (6.3)</td>
<td>17/197 (8.6)</td>
<td>30/404 (7.4)</td>
<td></td>
</tr>
<tr>
<td><strong>Living arrangements, n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lives by herself</td>
<td>12/207 (5.8)</td>
<td>14/197 (7.1)</td>
<td>26/404 (6.4)</td>
<td>.25</td>
</tr>
<tr>
<td>Lives with partner/spouse/children</td>
<td>155/207 (74.9)</td>
<td>157/197 (79.7)</td>
<td>312/404 (77.2)</td>
<td></td>
</tr>
<tr>
<td>Lives with other relatives/Friends</td>
<td>40/207 (19.3)</td>
<td>26/197 (13.2)</td>
<td>66/404 (16.3)</td>
<td></td>
</tr>
<tr>
<td><strong>Employment status, n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>25/207 (12.1)</td>
<td>39/197 (19.8)</td>
<td>64/404 (15.8)</td>
<td>.03</td>
</tr>
<tr>
<td>Not employed</td>
<td>182/207 (87.9)</td>
<td>158/197 (80.2)</td>
<td>340/404 (84.2)</td>
<td></td>
</tr>
<tr>
<td><strong>Home district, n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kisumu</td>
<td>60/207 (29.0)</td>
<td>67/197 (34.0)</td>
<td>127/404 (31.4)</td>
<td>.28</td>
</tr>
<tr>
<td>Others</td>
<td>147/207 (71.0)</td>
<td>130/197 (66.0)</td>
<td>277/404 (68.6)</td>
<td></td>
</tr>
<tr>
<td><strong>Pregnancy duration at recruitment, n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14-28 weeks</td>
<td>161/207 (77.8)</td>
<td>155/197 (78.7)</td>
<td>316/404 (78.2)</td>
<td>.83</td>
</tr>
<tr>
<td>29-36 weeks</td>
<td>46/207 (22.2)</td>
<td>42/197 (21.3)</td>
<td>88/404 (21.8)</td>
<td></td>
</tr>
<tr>
<td><strong>Total pregnancies, including current, median (IQR)</strong></td>
<td>3 (2, 4)</td>
<td>3 (2, 4)</td>
<td>3 (2, 4)</td>
<td>.38</td>
</tr>
<tr>
<td><strong>Total number of children</strong>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Living, mean (SD)</td>
<td>1.0 (1.5)</td>
<td>1.7 (1.0)</td>
<td>1.8 (1.3)</td>
<td>.045</td>
</tr>
<tr>
<td>Dead, mean (SD)</td>
<td>1.2 (1.6)</td>
<td>1.4 (1.3)</td>
<td>1.3 (1.5)</td>
<td>.34</td>
</tr>
<tr>
<td><strong>Duration of HIV+ status, n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 year or less</td>
<td>136/207 (65.7)</td>
<td>132/197 (67.0)</td>
<td>268/404 (66.3)</td>
<td>.63</td>
</tr>
<tr>
<td>2–4 years</td>
<td>49/207 (23.7)</td>
<td>40/197 (20.3)</td>
<td>89/404 (22.0)</td>
<td></td>
</tr>
<tr>
<td>5 or more years</td>
<td>22/207 (10.6)</td>
<td>25/197 (12.7)</td>
<td>47/404 (11.6)</td>
<td></td>
</tr>
<tr>
<td>Disclosed status to partner/spouse, n/N (%)</td>
<td>120/207 (58.0)</td>
<td>116/197 (58.9)</td>
<td>236/404 (58.4)</td>
<td></td>
</tr>
<tr>
<td><strong>Knows the HIV status of spouse/partner, n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>110/207 (53.1)</td>
<td>105/196 (53.6)</td>
<td>215/403 (53.4)</td>
<td>.93</td>
</tr>
<tr>
<td>Don’t know</td>
<td>97/207 (46.9)</td>
<td>91/196 (46.4)</td>
<td>188/403 (46.7)</td>
<td></td>
</tr>
<tr>
<td><strong>HIV status of spouse/partner, n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>77/110 (70.0)</td>
<td>76/105 (72.4)</td>
<td>153/215 (71.2)</td>
<td>.70</td>
</tr>
<tr>
<td>Negative</td>
<td>33/110 (30.0)</td>
<td>29/105 (27.6)</td>
<td>62/215 (28.8)</td>
<td></td>
</tr>
</tbody>
</table>

Continued
significance in the final model, except for general health perception and the study group. Participants with better health perceptions had a lower hazard of being LTFU (HR=0.99; 95% CI=0.98, 1.00). Being in the intervention group had a significantly lower hazard of being LTFU overall (HR=0.29; 95% CI=0.12, 0.69). However, a significant interaction term shows that the comparative hazard of being LTFU for intervention versus the control group was higher between 6 and 14 weeks postpartum (as reported earlier, the second largest number of participants LTFU were in the intervention group in the 6–14-week period).

**Effect of Cell Phone Counseling on EID Uptake**

Uptake of infant HIV testing was significantly higher among participants in the intervention arm compared with those in standard care (92.8% vs. 68.1%, respectively; \( P < .001 \)), followed until 14 weeks postpartum (Table 4). Stillbirths (n=9 in intervention; n=8 in control) and infant deaths (n=3 in intervention; n=1 in control) prior to 6 weeks were excluded from this analysis. PCR testing was conducted at a median of 44 days postpartum (interquartile range [IQR]=42, 49). The time of HIV testing of infants did not differ between

---

**TABLE 1. Continued**

<table>
<thead>
<tr>
<th></th>
<th>Intervention (n=207)</th>
<th>Control (n=197)</th>
<th>Total (N=404)</th>
<th>( P ) Value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ART use, n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Naive</td>
<td>118/207 (57.0)</td>
<td>113/197 (57.4)</td>
<td>231/404 (57.2)</td>
<td>.94</td>
</tr>
<tr>
<td>Experienced</td>
<td>89/207 (43.0)</td>
<td>84/197 (42.6)</td>
<td>173/404 (42.8)</td>
<td></td>
</tr>
<tr>
<td><strong>Type of ARV treatment assigned at PMTCT, n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Option A (only AZT)</td>
<td>84/207 (40.6)</td>
<td>84/197 (42.6)</td>
<td>168/404 (41.6)</td>
<td>.88</td>
</tr>
<tr>
<td>Option B (3 ARVs)</td>
<td>123/207 (59.4)</td>
<td>113/197 (57.4)</td>
<td>236/404 (58.4)</td>
<td></td>
</tr>
<tr>
<td><strong>CD4 cell counts at baseline,( b ) n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;350 cells/ml</td>
<td>50/162 (30.8)</td>
<td>53/156 (34.0)</td>
<td>103/318 (32.4)</td>
<td>.64</td>
</tr>
<tr>
<td>351–500 cells/ml</td>
<td>37/162 (22.8)</td>
<td>39/156 (25.0)</td>
<td>76/318 (23.9)</td>
<td></td>
</tr>
<tr>
<td>&gt;500 cells/ml</td>
<td>75/162 (46.3)</td>
<td>64/156 (41.0)</td>
<td>139/318 (43.7)</td>
<td></td>
</tr>
<tr>
<td><strong>Viral load at baseline,( c ) n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1000 copies/ml</td>
<td>70/90 (77.8)</td>
<td>60/72 (83.3)</td>
<td>130/162 (80.3)</td>
<td></td>
</tr>
<tr>
<td>&gt;1001 copies/ml</td>
<td>20/90 (22.2)</td>
<td>12/72 (16.7)</td>
<td>32/162 (19.9)</td>
<td></td>
</tr>
<tr>
<td><strong>Stigma score [16–64], n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low (16–40)</td>
<td>146/207 (70.5)</td>
<td>135/197 (68.5)</td>
<td>281/404 (69.6)</td>
<td>.32</td>
</tr>
<tr>
<td>Moderate (41–52)</td>
<td>59/207 (28.5)</td>
<td>62/197 (31.5)</td>
<td>121/404 (30.0)</td>
<td></td>
</tr>
<tr>
<td>High (53–64)</td>
<td>2/207 (1.0)</td>
<td>0/197 (0.0)</td>
<td>2/404 (0.5)</td>
<td></td>
</tr>
<tr>
<td><strong>Depression score [10–60], n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No depression (&lt;16)</td>
<td>136/207 (65.7)</td>
<td>107/197 (54.3)</td>
<td>243/404 (60.2)</td>
<td>.02</td>
</tr>
<tr>
<td>Sign of depression (≥16)</td>
<td>71/207 (34.3)</td>
<td>90/197 (45.7)</td>
<td>161/404 (39.9)</td>
<td></td>
</tr>
<tr>
<td><strong>General health perception [0–100], n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good/excellent (&gt;60)</td>
<td>99/207 (47.8)</td>
<td>76/197 (38.6)</td>
<td>175/404 (43.3)</td>
<td>.06</td>
</tr>
<tr>
<td>Average/poor/very poor (≤60)</td>
<td>108/207 (52.2)</td>
<td>121/197 (61.4)</td>
<td>229/404 (56.7)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: ART, antiretroviral therapy; ARV, antiretroviral; AZT, zidovudine; CD4, cluster of differentiation 4; IQR, interquartile range; PMTCT, prevention of mother-to-child transmission; SD, standard deviation.

<sup>a</sup> \( P \) value is from chi-square test for discrete variables; Mann-Whitney U test for median and \( t \) test in case of means for continuous variables.

\( b \) CD4 cell counts were obtained from clinic records; there are missing values due to missing records.

\( c \) Viral load tests were initiated late in the course of the study, thus results are available for only some participants.
the 2 groups. PCR test results were available for 308 infants. Of these, 9 infants (2.9%) tested positive, 7 infants in the intervention group and 2 infants in the control group (HIV incidence rate=0.39 per 100 infant-weeks; 95% CI=0.20, 0.75) (Table 5).

On binary logistic regression analysis, lower MPR, indicating poor adherence, was the main predictor of HIV infection among infants (Table 6). Overall, 20.8%, 28.4%, and 17.5% of the participants had MPR less than 90% at delivery, 6 weeks postpartum, and 14 weeks postpartum, respectively; there was no difference between the intervention and control groups (data not shown). Among the 9 infants who tested HIV positive, for 6 of the infants, the mother had MPR<90% at delivery and at 6 weeks postpartum.

**Uptake of Maternal Health Services**

The ANC attendance rates were higher among participants in the intervention arm than the control arm: 54.6% of the participants in the intervention arm completed more than 75% of the required number of visits compared with 41.8% in the control arm (P=.03) (Table 7). The required number of visits varied for participants depending on when in their pregnancy they registered for ANC attendance.

**FIGURE 3. Participant Retention in Care at Delivery and 6 Weeks and 14 Weeks Postpartum, Kisumu, Kenya**

Followed until delivery: intervention, 197/207; control, 153/197.

Followed until 6 weeks postpartum: intervention, 183/195; control, 137/188. Denominators exclude 17 participants total with still births (intervention, 9; control, 8) and 4 participants with infant deaths (intervention, 3; control, 1).

Followed until 14 weeks postpartum: intervention, 160/192; control, 125/188. Denominator for the intervention arm excludes 2 participants with infant deaths after 6 weeks postpartum.

---

**TABLE 2. Intervention Exposure and Retention Among Intervention Participants (n=207)**

<table>
<thead>
<tr>
<th>Percentage of Required Counseling Sessions Attended</th>
<th>Distribution of Participants Who Attended Calls (%)</th>
<th>Duration of Calls (minutes) Mean (SD)</th>
<th>Participants Retained Until 14 Weeks Postpartum (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤25%</td>
<td>10.7</td>
<td>18.7 (19.5)</td>
<td>55.6</td>
</tr>
<tr>
<td>26%-50%</td>
<td>17.0</td>
<td>10.5 (4.8)</td>
<td>63.3</td>
</tr>
<tr>
<td>51%-75%</td>
<td>35.0</td>
<td>7.0 (2.7)</td>
<td>92.7</td>
</tr>
<tr>
<td>&gt;75</td>
<td>37.4</td>
<td>7.9 (4.2)</td>
<td>89.3</td>
</tr>
</tbody>
</table>

Abbreviations: ANOVA, analysis of variance; SD, standard deviation.

a One way ANOVA. Differences between categories were significant at P<.001.

b Chi-square test. Differences between categories were significant at P<.001.

**: ANC attendance was higher in the intervention arm than the control arm.
TABLE 3. Predictors of Loss to Follow-Up at Delivery and 6 and 14 Weeks Postpartum (Complementary Log-Log Regression Model)

<table>
<thead>
<tr>
<th>Time point</th>
<th>Baseline Model (N=404)</th>
<th>Final Model (N=404)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>P Value</td>
</tr>
<tr>
<td>Delivery</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6 weeks postpartum</td>
<td>0.21 (0.10, 0.43)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>14 weeks postpartum</td>
<td>0.76 (0.49, 1.13)</td>
<td>.18</td>
</tr>
<tr>
<td>Living arrangements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lives with partner/spouse/children</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Lives by herself or with other relatives/friends</td>
<td>0.74 (0.29, 1.65)</td>
<td>.50</td>
</tr>
<tr>
<td>Employment status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not employed</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>1.69 (1.07, 2.33)</td>
<td>.03</td>
</tr>
<tr>
<td>Home district</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kisumu</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>1.22 (0.80, 1.73)</td>
<td>.35</td>
</tr>
<tr>
<td>Pregnancy duration at recruitment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14–28 weeks</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>29–36 weeks</td>
<td>1.08 (0.69, 1.58)</td>
<td>.73</td>
</tr>
<tr>
<td>Total pregnancies, including current</td>
<td>1.00 (0.90, 1.10)</td>
<td>.95</td>
</tr>
<tr>
<td>Has had HIV-affected child (living/died)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>0.89 (0.47, 1.56)</td>
<td>.72</td>
</tr>
<tr>
<td>Duration of knowing own HIV status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;5 years</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>≥5 years</td>
<td>0.62 (0.39, 0.97)</td>
<td>.04</td>
</tr>
<tr>
<td>Disclosure of HIV status to spouse/partner</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not disclosed</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Disclosed</td>
<td>0.68 (0.44, 1.00)</td>
<td>.05</td>
</tr>
<tr>
<td>Awareness of partner’s HIV status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does not know</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Knows, partner positive</td>
<td>0.62 (0.38, 0.96)</td>
<td>.03</td>
</tr>
<tr>
<td>Knows, partner negative</td>
<td>0.81 (0.45, 1.34)</td>
<td>.46</td>
</tr>
<tr>
<td>Stigma score (16–64) at baseline</td>
<td>1.01 (0.98, 1.04)</td>
<td>.57</td>
</tr>
<tr>
<td>Depression score (10–60) at baseline</td>
<td>1.02 (1.00, 1.03)</td>
<td>.046</td>
</tr>
<tr>
<td>General health perception (0–100) at baseline</td>
<td>0.98 (0.97, 0.99)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Experience of physical/sexual violence from partner at baseline</td>
<td>1</td>
<td>.15</td>
</tr>
</tbody>
</table>
ANC services. The relative risk ratio of not completing at least 50% of the visits was lower in the intervention arm (RRR=0.88; 95% CI=0.80, 0.98) compared with the control group.

For complete PNC coverage, the national program requires women who deliver to be seen within 24 to 48 hours after delivery followed by visits between 7 and 14 days postpartum and at 6 weeks postpartum. PNC attendance rates (3 PNC visits) were higher among participants in the intervention arm compared with the control (20% vs. 13%, respectively; RRR of non-completion of 3 PNC visits, intervention vs. control=0.93; 95% CI=0.86, 1.00) (Table 7). Attendance at the 6-week PNC visit that coincides with HIV testing of HIV-exposed infants was also higher in the intervention arm than the control (82% vs. 71%, respectively; RRR of non-attendance, intervention vs. control=0.89; 95% CI=0.82, 0.97).

### DISCUSSION

In this randomized study, we demonstrated that one-on-one individually tailored, theory-based counseling delivered via cell phone was highly effective in retaining mothers with HIV infection in care at delivery and at 6 and 14 weeks postpartum compared with standard care. About one-quarter of participants in the control group were LTFU in the period between recruitment and delivery. The number of dropouts reduced significantly after delivery indicating that the risk of LTFU is greatest prior to delivery. Most studies have examined retention in the postpartum period; non-retention in the antenatal period requires special attention. Retention also declined between 6 and 14 weeks postpartum (20 women moved out of Kisumu, 14 in the intervention arm and 6 in the control arm; 8 women were untraced, 3 in the intervention arm and 5 in the control). The women appeared to have waited for the HIV test

### TABLE 3. Continued

<table>
<thead>
<tr>
<th></th>
<th>Baseline Model (N=404)</th>
<th>Final Model (N=404)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>P Value</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HR (95% CI)</td>
</tr>
<tr>
<td>Baseline viral load</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1000</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>≥1000</td>
<td>1.02 (0.40, 1.86)</td>
<td>.96</td>
</tr>
<tr>
<td>CD4 count at enrollment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤350</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>351–500</td>
<td>0.70 (0.34, 1.33)</td>
<td>.30</td>
</tr>
<tr>
<td>&gt;500</td>
<td>0.81 (0.45, 1.36)</td>
<td>.47</td>
</tr>
<tr>
<td>ART experience</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Naive</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Experienced</td>
<td>0.80 (0.52, 1.17)</td>
<td>.26</td>
</tr>
<tr>
<td>Study group assignment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>0.22 (0.11, 0.46)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction between time point and study group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delivery</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6 weeks postpartum * Intervention</td>
<td>1.10 (0.21, 2.33)</td>
<td>.90</td>
</tr>
<tr>
<td>14 weeks postpartum * Intervention</td>
<td>3.35 (2.19, 2.41)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

**Abbreviations:** ART, antiretroviral therapy; CD4, cluster of differentiation 4; CI, confidence interval; HR, hazard ratio.

**Note:** Both baseline and final models are controlled for age, education, and marital status. All baseline models were tested for interaction effect with the time variable. The time interaction effect is included only if it is significant (P<.05). Final model comprised of variables that were significant in the baseline models.
results before moving out of Kisumu, suggesting that EID is a critical follow-up point in PMTCT programs.

The counseling intervention and positive health perceptions were independent predictors for participant retention at all 3 time points in the study. However, it is important for program managers to also focus on the other factors found to negatively influence retention in the baseline models controlled for age, education, marital status, and the intervention. These factors included the presence of depression and being employed, even though these factors were not significant in the multivariate model. A higher proportion of women in the control arm were employed and being employed was a predictor of loss to follow-up. Disclosure and knowledge of the HIV status of the spouse would have made it easier for women to receive counseling calls, talk freely, attend scheduled clinic visits, and get their infant tested for HIV infection on time. It is important for health workers to encourage and support disclosure and partner testing. We documented a trend toward a linear relationship in the effect of the intervention—retention increased with higher exposure to counseling—suggesting a benefit of the ongoing relationship between the women and their cell phone counselors. The counseling intervention was also highly effective in promoting the uptake of EID; we observed a significantly higher uptake of EID among participants in the intervention arm compared with the control arm.

With the widespread use of cell phones in Africa, other studies have evaluated the use of
cell phone technology for health messaging to promote retention in care, but almost all studies have used SMS. This is the first study to provide focused and individualized counseling delivered via phone technology from one offsite central location. We report much higher retention rates at 6 weeks than a 2-arm randomized controlled study by Odeny et al. (2014) that used SMS to increase attendance at maternal postpartum clinic (intervention vs. control: 19.6% vs. 11.8%; relative risk=1.66; 95% CI=1.02, 2.70) in the Nyanza region of Kenya. A second, more recent randomized cluster study, conducted in Homa Bay, Nyanza, by Kassaye et al. (2016), used 2-way SMS between counselors and patients to improve retention, uptake of EID, and face-to-face communication. The authors report high retention at 6 weeks in both the intervention arm (87%; 244/280) and the control arm (84%; 227/270), and similarly improved communication between patients and counselors in both arms. Both studies report very high rates for infant testing (Odeny et al.: intervention vs. control, 92% vs. 85%; Kassaye et al.: intervention vs. control, 88% vs. 89%) and low HIV positivity among infants tested (Odeny et al.: 1.5%; Kassaye et al.: 0.9%). In contrast, we report significantly higher HIV testing in the intervention arm (94%) compared with standard care (68%), but also higher positivity rates with more HIV PCR-positive infants in the intervention arm than the control arm (2.9%; n=7 in the intervention and n=2 in the control). Differential follow-up rates in the 2 arms in our study may have contributed to these results—that is, a higher proportion of participants in the control arm were not retained and their infants were not tested. It is also possible that there may be programmatic differences in the 2 geographic areas, Homa Bay and Kisumu, where the studies have been conducted. Both are located in the same province on the shores of Lake Victoria, but Homa Bay is a high HIV prevalence area with a long-standing, mature HIV program with several programmatic interventions in place and a population familiar with HIV. Although there was no significant difference between the 2 groups of participants with regard to MPR, we found lower MPR to be an independent predictor of HIV

<table>
<thead>
<tr>
<th>TABLE 6. Determinants of HIV Infection Among Exposed Infants (Binary Logistic Regression Model)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ART experience</td>
</tr>
<tr>
<td>Naive</td>
</tr>
<tr>
<td>Experienced</td>
</tr>
<tr>
<td>Pregnancy duration at recruitment</td>
</tr>
<tr>
<td>14–28 weeks</td>
</tr>
<tr>
<td>29–36 weeks</td>
</tr>
<tr>
<td>Place of delivery</td>
</tr>
<tr>
<td>Home</td>
</tr>
<tr>
<td>Institutional</td>
</tr>
<tr>
<td>Infant feeding</td>
</tr>
<tr>
<td>Mixed or complementary feeding</td>
</tr>
<tr>
<td>Exclusive breastfeeding</td>
</tr>
<tr>
<td>Medication Possession Ratio</td>
</tr>
<tr>
<td>≤ 90% or less</td>
</tr>
<tr>
<td>&gt; 90%</td>
</tr>
<tr>
<td>Study group assignment</td>
</tr>
<tr>
<td>Control</td>
</tr>
<tr>
<td>Intervention</td>
</tr>
</tbody>
</table>

Abbreviations: ART, antiretroviral therapy; CI, confidence interval; OR, odds ratio.
Note: Regression model is adjusted for mother’s age and education.
infection among infants. Thus, while retention and uptake of EID improved with our intervention, adherence remained a concern, especially in about one-fifth of the participants. Our findings highlight the need to closely follow pregnant women living with HIV in PMTCT programs, especially during the antenatal period and after EID (around the 6-week postnatal visit) with an increased focus on collection of medication and adherence to treatment. HIV transmission continues to remain a serious concern.

The counseling intervention resulted in higher antenatal and postnatal attendance rates in the intervention group compared with the control group. Interestingly, our postnatal attendance rates at 6 weeks (intervention vs. control, 80.1% vs. 70.5%) were higher than those reported by the 2014 Kenya Demographic and Health Survey for the general population of women in Nyanza and Kenya (36.7% women in Nyanza and 43.0% in Kenya had no PNC visit at all). Similarly, institutional delivery rates were similar across the 2 groups (average 94.4%) but significantly higher than institutional delivery rates reported by the Demographic and Health Survey among the general population of women in Kisumu (69.5%) and Kenya (61.2%). It appears that mothers living with HIV who are enrolled in the PMTCT program are more likely to obtain maternal health care services than the general population.

Although cell phone coverage has increased dramatically across many countries in the continent and is reported to be around 80% in Kenya, many women (355 women) did not have access to cell phones and thus could not participate in the study and benefit from the intervention. Further, counselors found it challenging to reach participants and make calls, taking on average more than 4 attempts to make a successful call. Finding a suitable window of time for the client to be able to talk freely and having access to phones shared with a spouse or family member are substantial barriers. As the intervention was found to be very effective in retaining patients in care and promoting uptake of EID, programs implementing this intervention may consider the provision of cell phones to women who are at higher risk of loss to follow-up and do not have access to a phone. The intervention tested is resource intensive, which may be a concern for programs considering scale up. Programs could consider reducing the total number of calls by about 30% based on the data that shows that the highest retention was observed among participants who attended between 51% and 75% of the calls. Alternatively, programs may consider allocating the intervention only to participants at higher risk of loss to follow-up, such as those reporting depressive symptoms or those who may not have disclosed their HIV status to their

### TABLE 7. Uptake of Maternal and Child Health Services and Infant HIV PCR testing

<table>
<thead>
<tr>
<th>Attended³:</th>
<th>Intervention n/N (%)</th>
<th>Control n/N (%)</th>
<th>P Value⁴</th>
<th>Unadjusted RRR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50% or less of required ANC visits</td>
<td>60/187 (32.1)</td>
<td>81/182 (44.5)</td>
<td>0.88 (0.80, 0.98)</td>
<td></td>
</tr>
<tr>
<td>51%–75% of required ANC visits</td>
<td>25/187 (13.4)</td>
<td>25/182 (13.7)</td>
<td>.03</td>
<td></td>
</tr>
<tr>
<td>76%–100% of required ANC visits</td>
<td>102/187 (54.6)</td>
<td>76/182 (41.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Received complete ANC package⁵</td>
<td>192/207 (92.8)</td>
<td>181/196 (92.4)</td>
<td>.88</td>
<td>1.00 (0.95, 1.05)</td>
</tr>
<tr>
<td>Attended 3 PNC visits per national protocol</td>
<td>42/207 (20.3)</td>
<td>25/197 (12.7)</td>
<td>.04</td>
<td>0.93 (0.86, 1.00)</td>
</tr>
<tr>
<td>Attended 6-week PNC visit</td>
<td>170/207 (82.1)</td>
<td>139/197 (70.6)</td>
<td>.006</td>
<td>0.89 (0.82, 0.97)</td>
</tr>
<tr>
<td>Delivered at a health facility⁴</td>
<td>188/197 (95.4)</td>
<td>143/153 (93.5)</td>
<td>.42</td>
<td>0.98 (0.93, 1.03)</td>
</tr>
<tr>
<td>Infants with full primary immunization⁶</td>
<td>156/160 (97.5)</td>
<td>121/125 (96.8)</td>
<td>.72</td>
<td>1.01 (0.97, 1.05)</td>
</tr>
</tbody>
</table>

Abbreviations: ANC, antenatal care; BCG, bacille Calmette-Guérin; CI, confidence interval; DPT, diphtheria, pertussis, and tetanus; PNC, postnatal care; RRR, relative risk ratio.

³ Chi-square test.
⁴ Assessed among those who required at least 1 ANC visit from the time of recruitment until delivery.
⁵ ANC package was considered complete if the following were done: hemoglobin, venereal disease research labs, blood group, and at least 1 urine test.
⁶ Assessed among those who were followed until delivery.

Full primary immunization was considered if the infant received BCG + polio at birth, and DPT + polio at 6, 10, and 14 weeks of age.
partners/families. Rigorously monitored program data can then be used to assess impact at scale.

Limitations
The study is not without limitations. We observed a very high rate of refusal during the screening process (564 refusals and 333 women who postponed recruitment and did not return), indicating a reluctance to be identified and contacted regularly, suggesting that stigma is still deeply entrenched in the community. Other studies have also reported persistent HIV-related stigma in African communities and the role of stigma in lower uptake of PMTCT services and EID in Kenya.\textsuperscript{19–31} Efforts must be made to reduce stigma in community. It is of note, however, that among study participants, stigma did not have an effect on retention or uptake of EID. As the intervention moves from a research setting to program delivery when scaled up, a high refusal rate has the potential to attenuate the population-level impact of the intervention. A detailed discussion of the program and its benefits (personalized and confidential access to counselors, information and support) by clinic staff may help to overcome stigma and increase participation, especially among women who are vulnerable and more likely to be lost to follow-up. Further, the lack of access to cell phones limited inclusion of a large number of women, which could have biased the sample. Therefore, the results should be interpreted within the given context. In our assessment of retention, we excluded women who had a stillbirth or infant death prior to 6 and 14 weeks postpartum, as these women would not need to attend child care services. This could have led to an underestimation of the infant HIV positivity rate as the stillbirths or infant deaths could have been due to in utero HIV infection. Lastly, we used MPR as a measure of adherence; it is, however, important to note that col- infection. Lastly, we used MPR as a measure of adherence; it is, however, important to note that col-
i-...ects. Further, the lack of access to cell phones limited inclusion of a large number of women, which could have biased the sample. Therefore, the results should be interpreted within the given context. In our assessment of retention, we excluded women who had a stillbirth or infant death prior to 6 and 14 weeks postpartum, as these women would not need to attend child care services. This could have led to an underestimation of the infant HIV positivity rate as the stillbirths or infant deaths could have been due to in utero HIV infection. Lastly, we used MPR as a measure of adherence; it is, however, important to note that collection of medications does not confirm that participants consumed the medications. Future studies should include a more direct measure of adherence such as therapeutic drug monitoring or ante- and postnatal viral load monitoring.

CONCLUSION
In conclusion, the one-on-one individually tailored, theory-based counseling delivered via cell phone was very effective in retaining mothers with HIV in care and in promoting the uptake of EID and antenatal and postnatal care services. Within the intervention, a greater emphasis is required on the collection of medications and adherence.

Acknowledgments: The study team wishes to thank the following people for their hard work on this study: health facility staff Jacob Ngachra (Ober Kamloth), Beatrice Obare (Airport), Rachel Otunge (Ololla), Jane Ambiya (Simba Opepo), Zebedi Ambiya (Disciples of Mercy), Kaito Obiet (Joneses Ring Road), Celestino Okello (Nyawenda), Alice Sagwe (Kowino Health Center), Hilda Omond (K-meli), Hezbain Mwakachala (Marie Stapes International Kenya), Cleophas Sore (Migosi), Clare Gichana (St. Elizabeth Chiga), Martin Agunda (Chiga Health Center), and Grace Mirongo (Gitaa); county health officials Eunice Kinywa (DASCO), Dr. Charles Okal (CASCO), and Samuel Ochieng (Lab Manager); Kemri Alupe Lab staff Professor Mathilu Mwau, Joy Mwende Ndunda, and Catherine Akinyi; data collection team including research assistants Kevin Oniala, Florence Anyango, Christine Mbia, Nelly Adanjo Adhiambo, Franklin Otieno, Beatrice Ogaga, and Joshua Okoth; counselors including Georgina Oyende, Rosecloay Milanya, Mercyline A. Opande, Ruth Akinyi Muganda, and Alice Ojwang; coordinator Danmark Owuor; information technology Misra Ashutosh and Isakk Adan; and study coordinator Janet Munyasya.

Funding: This study was made possible through support provided by the U.S. President’s Emergency Plan for AIDS Relief and the U.S. Agency for International Development (USAID) via HVICore, a Task Order funded by USAID under the Project SEARCH indefinite quantity contract (contract no. AID-OAA-TO-11-00060). The Task Order is led by the Population Council in partnership with the Elizabeth Glaser Pediatric AIDS Foundation, Palladium and the University of Washington. The content of this article is the responsibility of the authors and does not necessarily reflect the views of USAID or the U.S. government.

Competing Interests: None declared.

REFERENCES


Peer Reviewed

Received: July 2, 2018; Accepted: March 12, 2019; First Published Online: May 29, 2019


© Sarna 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-18-00241
“It’s Not Like Taking Chocolates”: Factors Influencing the Feasibility and Sustainability of Universal Test and Treat in Correctional Health Systems in Zambia and South Africa

Stephanie M. Topp,a,b Candice M. Chetty-Makkan,c Helene J. Smith,b Lucy Chimoyi,c Christopher J. Hoffmann,d Katherine Fielding,e Stewart E. Reid,b,f Abraham J. Olivier,g Harry Hausler,g Michael E. Herce,b,h Salome Charalambousc,i

Universal test and treat may be feasible even in highly resource-constrained correctional facilities. Sustainability and impact of such services require a supportive policy environment, robust service delivery systems, adequate resourcing, and close attention to the psychosocial factors influencing incarcerated persons’ willingness to engage in HIV treatment.

ABSTRACT

Background: Sub-Saharan African correctional facilities concentrate large numbers of people who are living with HIV or at risk for HIV infection. Universal test and treat (UTT) is widely recognized as a promising approach to improve the health of individuals and a population health strategy to reduce new HIV infections. In this study, we explored the feasibility and sustainability of implementing UTT in correctional facilities in Zambia and South Africa.

Methods: Nested within a UTT implementation research study, our qualitative evaluation of feasibility and sustainability used a case-comparison design based on data from 1 Zambian and 3 South African correctional facilities. Primary data from in-depth interviews with incarcerated individuals, correctional managers, health care providers, and policy makers were supplemented by public policy documents, study documentation, and implementation memos in both countries. Thematic analysis was informed by an empirically established conceptual framework for health system analysis.

Results: Despite different institutional profiles, we were able to successfully introduce UTT in the South Africa and Zambian correctional facilities participating in the study. A supportive policy backdrop was important to UTT implementation and establishment in both countries. However, sustainability of UTT, defined as relevant government departments’ capacity to independently plan, resource, and administer quality UTT, differed. South Africa’s correctional facilities had existing systems to deliver and monitor chronic HIV care and treatment, forming a “scaffolding” for sustained UTT despite some human resources shortages and poorly integrated health information systems. Notwithstanding recent improvements, Zambia’s correctional health system demonstrated insufficient material and technical capacity to independently deliver quality UTT. In the correctional facilities of both countries, inmate population dynamics and their impact on HIV-related stigma were important factors in UTT service uptake.

Conclusion: Findings demonstrate the critical role of policy directives, health service delivery systems, adequate resourcing, and population dynamics on the feasibility and likely sustainability of UTT in corrections in Zambia and South Africa.

BACKGROUND

Sub-Saharan African correctional facilities concentrate large numbers of people living with HIV and at risk for HIV infection. They include people from impoverished communities, those with poor baseline health, and those who engage in high-risk behaviors, such as commercial sex workers or individuals with substance use disorders.1,2 As a result, incarcerated populations in Sub-Saharan African correctional facilities, who encompass both sentenced offenders and those detained and awaiting trial (on remand), experience higher HIV and tuberculosis (TB) prevalence than the general population.3
Studies from Zambia and South Africa, 2 countries among those having the largest correctional populations and generalized HIV epidemics in Southern Africa, have reported HIV prevalence among incarcerated persons between 12.5% and 27.4% (Zambia) and between 7.2% and 18.9% (South Africa), and TB prevalence of 0.34% to 7.6% (Zambia) and 0.71% to 3.6% (South Africa). The right to health obligates governments to provide health care to incarcerated people. To effectively prevent, treat, and care for HIV among incarcerated people, services must be tailored to their unique needs and HIV risk profile. Such efforts are necessary both because the demographic characteristics of and disease burden faced by incarcerated populations differ from those of the general population in the community and because the operational environment of correctional facilities are particular.

Universal test and treat (UTT) is widely recognized as a promising approach to improve the health of individuals living with HIV and a population health strategy to reduce the incidence of new HIV infection. The benefits of early treatment have been demonstrated in various clinical trials documenting reduced risk of HIV transmission when viral load is suppressed, and reductions in HIV incidence with increasing population-level antiretroviral therapy (ART) coverage. Proponents of UTT have also highlighted the potential for a scaled-up approach to reduce the many barriers faced by people living with HIV to starting ART, and World Health Organization guidelines now recommend immediate ART initiation for all people living with HIV globally. Yet little is known about barriers to or facilitators of establishing and maintaining UTT in nontraditional programmatic settings such as correctional facilities.

With the aim of delivering a UTT intervention to incarcerated individuals in correctional health systems, we conducted the Treatment as Prevention (TasP) study in 3 correctional facilities in South Africa and 1 correctional facility in Zambia. Given the lack of evidence to date regarding UTT in correctional systems anywhere, a defined aim at the outset was to qualitatively explore the feasibility and compare the challenges to sustainability of this intervention within correctional facilities in 2 correctional systems.

**METHODS**

**Setting**

Zambia has a network of 88 correctional facilities (some male only, some co-joined male/female) including several large maximum and medium security sites and a host of smaller district and farm prisons with a total official capacity of 6,100. Occupancy levels are high. In 2018, the Zambia prison population was estimated to be 25,000, representing a 34% increase since 2015 and an occupancy rate of 300%. The facility in which the current study was located is a medium-security facility in Zambia’s capital city, Lusaka, and houses both convicted and “on-remand” individuals. It has a co-joined male and female wing, and at the time of study, it housed approximately 1,241 male and 129 female incarcerated persons, indicating an occupancy rate of over 300%. Administrative responsibility for corrections falls to the Zambia Correctional Service (ZCS) within the Ministry of Home Affairs.

South Africa has a network for some 243 correctional facilities ranging from maximum security to Community Corrections sites. South Africa’s correctional system serves a population of over 158,000 incarcerated people. It has the 12th highest incarcerated population in the world, with an occupancy rate of 135%. Correctional facilities are administrated by a dedicated government ministry, the Department of Correctional Services. The South African sites in this study included 3 correctional facilities with 8 units. The Johannesburg Correctional Facility included 1 male maximum, 1 male medium, and 1 female unit, for a total population of 9,171. In Breede Valley, Brakvallei Correctional Facility had 1 male maximum, 1 male medium, and 1 youth (ages 18–22) unit, and Worcester Correctional Facility had 1 female and 1 male maximum unit, for a combined total population of 4,173.

**Objectives**

In this nested qualitative study, we aimed to (1) critically examine and compare feasibility of establishing UTT within the Zambian and South African correction health systems, and (2) draw out and compare aspects of both the intervention and the different correctional health system contexts likely to support or inhibit sustainability of UTT in the future.

**Conceptual Framework**

This study was designed with the understanding that health systems are complex, meaning their function and performance are determined by the dynamic interactions between health system “hardware” and health system “software,” respectively. Drawing on the work of Sheikh and
colleagues, we define health system hardware as the tangible, material components of a system, such as infrastructure, health workers, drugs, and commodities. Health system software comprises the interests, values, relationships, and practice-based norms of the human stakeholders whose decisions and actions bring the health system to life. Critically, in health system analyses, recognition of the values, beliefs, and relationships of service users (clients and/or patients) is considered just as essential to understanding health system performance as those of providers, managers, and policy makers.

Feasibility studies typically seek to answer the question “Can this [intervention] be done in a given setting or context?” Hardware and software factors commonly considered in relation to health service feasibility include the willingness of patients, providers, and health planners and managers to participate; the perceived appropriateness and convenience of the intervention; availability of appropriate resourcing; and logistical systems required to support the intervention.

Sustainability is conceptually distinct from feasibility and, simply defined, may be thought of as the “capability of being maintained at a certain rate or level.” In the context of evaluating health programs such as UTT, sustainability may be considered:

> the ability . . . to function effectively, for the foreseeable future, with high treatment coverage, integrated into available health care services, with strong community ownership, using resources mobilized by the community and government.

Mirroring the inclusive understanding of health systems described above, evaluation of sustainability involves attention to the broader organizational and systems dynamics, not just the obvious issues of proximate resourcing or training considerations. Interactions between stakeholders, institutions, and beneficiaries (“relational” components) are particularly important in ensuring continuation of a program, including via planned or spontaneous adaptations. Our understanding of sustainability was informed by these considerations as well as Schell and colleagues’ domains of sustainability, which include strategic planning, organizational capacity, program adaptation, program evaluation, and communication, as well as political support, funding stability, partnerships, and public health impacts.

### Primary Data Collection

Nested within the intervention study, our qualitative evaluation of UTT feasibility and sustainability used a case-comparison design that considered Zambia’s and South Africa’s correctional health systems as distinct cases. Within each case, we sought primary data from in-depth interviews with male and female incarcerated respondents with and without HIV and corrections staff and policy makers. We conducted interviews with all respondent types at each site. The rationale for interviewing respondents without HIV who were not directly involved in the UTT intervention was to ensure data on the nature of correctional life and correctional health services more generally and obtain reflections on any impact UTT may have had on those services.

Recruitment for all interviews with incarcerated individuals was a mixture of purposive and opportunistic, targeting male and females with and without HIV at each site. Respondents with HIV were identified through the intervention study (i.e., during enrollment or treatment procedures); respondents without HIV were identified with the assistance of correctional health staff, but no participants were selected by correctional staff. All individuals were invited to participate by a member of the study team during private visits to the clinic.

Recruitment of correctional staff (both health and non-health officials) was a mixture of purposive and opportunistic, and it was based on the study team’s progressive identification of individuals involved in some aspect of health service delivery or planning within each correctional facility. Participation of any staff member typically had to be cleared through the facility in-charge, and recruitment was subsequently carried out via private invitation, in person, by a member of the study team. Selection of senior administrators and policy makers was wholly purposive and based on the investigators’ prior (and emerging) knowledge of their involvement and expertise in the study area. Recruitment was achieved via phone invitation, email, or occasionally in person.

All interviews were conducted in person by trained research assistants recruited separately in each country. All but one interviewer (in South Africa) had previous experience conducting qualitative research, and all interviewers received a 5-day training on interview guides and human subjects protection in corrections. Interviews ranged from 20 to 100 minutes, and all were audio-recorded with written consent from the
participants. Recordings were translated into English (as necessary) and transcribed in one step. Research assistants fluent in the language of the interview compared the transcript to the audio-recording to assess accuracy, completeness, and compliance with formatting requirements. Any anomalies were addressed by the interviewer or supervisor.

The Table summarizes the interviews conducted in Zambia (13 incarcerated individuals, 8 providers and administrators) and South Africa (37 incarcerated individuals, 13 providers and administrators). Interviews were guided by a common question guide developed for each of the 3 categories of respondents (incarcerated individuals with and without HIV, health care providers, corrections officials and policy makers). For each category of respondent, a common English question guide was developed during a multiphase consultation with Zambian and South African team members, with careful consideration given to cross-site differences (e.g., site-specific probes to capture local cultural and administrative context). The guide for incarcerated individuals focused on experiences with project (TasP) services as relevant and other experiences with the correctional health services, probing to understand matters of access, provider attitude, responsiveness, and overall acceptability. Health worker and correctional staff guides were framed by enquiries about the planning, communication, adaptation, resourcing, and perceived impact of UTT alongside questions regarding the organizational capacity and the day-to-day decisions faced by health and correctional personnel. Interviews with administrators and policy makers focused on higher-level concerns including the alignment of UTT with existing policy and programmatic priorities at the national and departmental/ministry levels, and the role and sources of funding and interinstitutional relationships. The English guide for the incarcerated respondents was translated into other national languages in Zambia (Nyanja, Bemba, and Tonga) and South Africa (Afrikaans, isiZulu, isiXhosa, Sesotho, and Setswana), back-translated, and pilot tested by the in-country teams. Health worker and policy maker interviews in both countries were conducted exclusively in English.

### Document and Program Data

In order to contextualize and triangulate our interview data, we additionally drew on publicly available national policy documents, study documentation, and investigator experience (e.g. implementation memos) to cross-reference and strengthen understanding and interpretation.

### Analysis

All transcripts were transcribed and translated (as necessary) in a single step into English by an experienced, bilingual research assistant. At both sites, transcripts were subsequently quality checked by a second bilingual investigator against the original recording.

Following close reading of a selection of the English transcripts, 3 authors (ST, HS, CC) consulted and co-developed a coding framework based initially on deductive reasoning (informed by Sheikh and colleagues’ framework15 and drawing on domains of feasibility and sustainability identified in the literature18,22–24) and refined inductively during the first several rounds of reading and coding. Transcripts were imported into Nvivo QSR (V.10 Australia) for coding and theming. For each site, ST plus a site-specific author co-coded several transcripts to validate and refine the codes and then coded 5 transcripts each to ensure consistency with high intercoder agreement. All remaining transcripts from Zambia and South Africa, respectively, were then coded independently.

### RESULTS

The qualitative findings of this study are presented in 2 sections. Built from interview data and investigator memos, Box 1 (Zambia) and Box 2 (South Africa) outline the living conditions and health service access of incarcerated respondents as

<table>
<thead>
<tr>
<th>TABLE. Interviews Conducted, by Site and Respondent Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>Zambia</td>
</tr>
<tr>
<td>South Africa, Johannesburg</td>
</tr>
<tr>
<td>South Africa, Cape Town</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>
BOX 1. Living Conditions and Health Service Access in the Zambian Facility

General Conditions
- Poor living conditions in most correctional facilities in the country, including the study site.
- Extreme overcrowding and lack of appropriate sleeping quarters.
- Male interviewees noted that a substantial proportion of male incarcerated persons lived in cells so crowded that they had to sleep in a seated, upright position on the floor due to lack of space and bedding.
- Most cells did not have an internal bathroom.
- Personal hygiene was hard to maintain due to the limited hours with access to running water or bathing trenches, and the high demand for the same.
- Female quarters were generally acknowledged to be less crowded, but still poor.
- Personal hygiene was reported as poor, and resources for cleaning products lacking.
- In both men and women’s facilities, incarcerated people described persistent concerns about communicable diseases, particularly tuberculosis.
- Poor food provided by Corrections—inadequate quantities and very poor quality—were a ubiquitous complaint. Most incarcerated people noted that without support from family or friends, they would not have food adequate to meet their nutritional needs.

Health Service Access
- Health service access required seeking permission from “senior inmates” and officers.
- During the day, male incarcerated individuals had to first report to their cell captains (those with delegated authority from officers) and ask to be placed on the official “sick list,” which gave them permission to go to the clinic. Once at the clinic, however, access to services was not guaranteed because opening hours were limited due to limited health workers (typically 1 clinical officer and/or nurse per day) being available and gatekeeping by peer educators (appointed incarcerated individuals) who managed clinic queues.
- At night, access to health care was more precarious, requiring the cell captain to bang on the door to attract the attention of the night guard, who then had to be convinced to contact a health worker.
- Incarcerated study participants noted that the attitudes and responsiveness varied substantially between health workers and security personnel. They also noted that access was particularly difficult in cases in which there was a lack of external (observable) symptoms.

Zambia Feasibility of UTT in Zambia
UTT services in the Zambian correctional facility were implemented as a predominantly parallel service, with components that were physically distinct from the correctional facility’s day-to-day health service. The decision to deliver UTT in this way was based on the project team’s a-priori knowledge of and pragmatic considerations relating to the constrained human and material resources available for health care within the Zambian correctional system. Standalone baseline clinical and laboratory evaluation and ART initiation services were delivered within the facility’s (internal) clinic, but they were staffed and resourced exclusively by the study. Where possible, some service components were integrated within existing correctional health processes. For example, HIV testing for the study was largely driven by counselors and peer educators who worked as part of the correctional health service for HIV testing and counseling. Overall, implementation of UTT in this manner proved highly feasible at the study site (a large, centrally located urban study facility), with interview data and investigator experience pointing to 3 major contributing factors as described below.

Alignment With National Policy
Interviews with senior correctional and Ministry of Health (MOH) officials demonstrated a high level of support for the idea of establishing and scaling up UTT to all Zambian correctional facilities. Such support was set against Zambia’s national policy of test and treat for HIV, introduced in 2017, which mirrors widespread political and bureaucratic enthusiasm for improved access to HIV care and treatment. One senior ZCS official noted that the policy meant that the opportunity to trial UTT in all correctional facilities was seen as an
important development and was in line with “mainstream health services.”

**“Open-Door” Policy and Long-Term Partnerships**

The Zambian Correctional Service has an open-door policy, working with trusted partners to strengthen facility-based health services. This policy operates in recognition of profound resourcing and capacity limitations within ZCS and enables correctional officers to work with nongovernment groups to address the most urgent health resourcing and health service gaps. Since 2009 when the unofficial policy was put in place, relationships with several long-standing partners have evolved.

The correctional service has opened its doors, to the stakeholders, to help provide health service, provide technical skills, provide funding, and this is helping to reorganize I think, the system for health delivery in the system. [ZCS official]

The existing partnership between the Zambian implementing NGO and ZCS was an important facilitator in the establishment of UTT in the Zambian site. The relationship and prior experiences underpinned both initial planning and the permissions provided by the ZCS Commissioner General and the subsequent routine access afforded to study staff by the facility manager. This relationship was a vital enabler. Ad hoc
challenges and miscommunications between study and correctional staff, an anticipatable challenge in any project, were substantially easier to manage in the context of the established relationship between the NGO and ZCS.

**Positive Attitudes Toward HIV Care and Treatment**

A final component of the feasibility of UTT in Zambia was the acceptable nature and comparatively strong demand for HIV testing and treatment services among the incarcerated population. Incarcerated individuals, health care providers, and ZCS officers all described the environment of the Zambian correctional facility as one that was generally open about and supportive of individuals seeking care. Several participants described how the encouragement they received to do an HIV test and look after their health when they first entered corrections were instrumental in their decisions to seek care.

> Even HIV, I did not even want to test when I was outside prison; because if I was outside prison maybe I was going to be saying I was bewitched. So I appreciate a lot because my coming to prison helped me to even test. [Female inmate, Zambia]

Demand for health services in the study facility was in part underpinned by a comparatively low-stigma environment and a notable culture of peer support in relation to HIV testing and treatment uptake.

> At the moment, stigma is no longer there, we just live as friends because everyone inside has a group. But there is no choosing like this one is taking ARVs or is on TB [treatment]. [Male with HIV, Lusaka]

Several respondents contrasted the comparatively lower HIV-related stigma in the men’s facility with the still prevalent concerns about stigma in the mainstream (noncorrections) community. Longer-term incarcerated respondents and corrections staff attributed this difference to a series of rolling education and sensitization programs carried out by various nongovernment groups over several years leading up to UTT. Such an environment was instrumental in both the strong early uptake and sustained demand for testing and treatment services delivered in the male section of the Zambian site of this project. Notably, however, female respondents in the much smaller female section described more concerns regarding HIV-related gossip and stigma.

> Sometimes you find an old [timer] inmate maybe stigmatizing another one. Mnhh, it was terrible, it was terrible. When your friend is very sick, looking at the congestion, you know we have to sleep like bumper to bumper. People were scared to sleep next to a sick person, thinking they will also catch that same disease. [Female without HIV, Lusaka]

**Sustainability of UTT in Zambian Correctional Facilities**

While the data point to the feasibility of implementing UTT as a largely vertical, NGO-supported service, we found more equivocal evidence in relation to the long-term sustainability of UTT, at least in terms of a government (ZCS and MOH)-led program. Mirroring some known issues, we identified the follow factors undermining the sustainability of the UTT approach: weak ZCS funding and resourcing for health activities; still limited organizational capacity of the Zambian correctional health system; and a policy backdrop that while supportive, did not make the need for UTT in correctional facilities sufficiently explicit. Reflecting both deductive and inductive themes, we identified barriers and facilitators relating to 3 major sustainability domains as outlined below.

**Policy Backdrop:** Zambia’s national policy for UTT introduced in 2017 helped to establish a backdrop of robust political support. However, ZCS and MOH stakeholders flagged the lack of explicit mention of correctional facilities or incarcerated populations as part of UTT scale-up as a potential problem. Lack of explicit mention of correctional services raised questions about whether national (MOH-controlled) resourcing for and support of UTT scale-up would cover the introduction and maintenance of UTT in correctional facilities, as an MOH official described:

> So the universal testing, counseling, and treatment for Zambia, it refers to the health facilities both public and private. [And although] that can be broadly translated to the prison service [at the moment] technically it does not apply; it does not apply to non-health settings. Which is a challenge. [MOH official]

In the context of the lack of depth in the health service leadership of ZCS, moreover, the absence of explicit guidelines to support the scale-up of UTT services was described as a considerable barrier.

**Funding and Resourcing:** The potentially narrow interpretation of “universal” test and treat as only applying to mainstream health services was a significant issue from both a resourcing and organizational standpoint. During the study, implementation of UTT in the Zambian correctional...
The feasibility of delivering UTT within South African correctional facilities drew on 2 major contributing factors.

Facility was dependent on external (study) funding. At the time of writing, additional external resources from 3 different donor-funded programs had been found to support continued provision of UTT in the short to mid term. But few immediate prospects for additional internal resources to maintain the staff or systems exist. Various stakeholders described a range of related resourcing concerns relating to the sustainability of UTT within the correctional service. These most notably included inadequate numbers and inappropriate skill mix of current human resources for health and largely inadequate health infrastructure that mitigated against the delivery of high-quality care.

You find that inmates will be many and there will only be one nurse. And then they will just say: “These 15 will be seen, the other ones will be seen in the afternoon.” But maybe they don’t attend to even those. [Male with HIV, Lusaka]

The infrastructure in correctional services for health . . . they are all dilapidated. And so we do not have an appropriate system where the services can be provided in a quality manner . . . . [ZCS official]

Organizational Capacity and Organizational Culture: Related to, but distinct from lack of resourcing as a barrier to sustained UTT, was the lack of depth and capacity for planning or delivery of care. While the aforementioned absence of human resource capacity and a dedicated health budget underpin these problems, lack of depth and capacity in leadership and supervisory roles also contributed to poorly constituted and weakly integrated health information systems and weak supply chains.

The other weakness is in terms of reporting systems. We have developed [paper-based] tools . . . but if we had computers, emails, so that we are able to monitor what is happening, in a second you are able to see what is happening in [X facility]. But we normally receive hard copies, when they send it from [X facility] to here. It may take 2 weeks to a month to reach here. [ZCS official]

Despite efforts to ameliorate such issues, lack of depth in health leadership also contributed to an often unresponsive health service culture, constituting an additional and distinct barrier to sustained provision of UTT in this setting. Various incarcerated respondents and a corrections-based health worker commented on the way these issues affected both HIV and more generalized health service access and uptake:

In the Corrections clinic you find that today the clinical officer is not there, today the nurse is not there. [Male with HIV, Lusaka]

You know they don’t treat us as patients. [Corrections health providers] treat us as criminals . . . so even if you come for treatment when you are sick, they are harsh on you. [Male with HIV, Lusaka]

South Africa
Feasibility of UTT in South Africa

Implementation of UTT in the 3 South African correctional facilities was negotiated between local project teams and Department of Correctional Services (DCS) staff separately at each of the 3 sites. In Johannesburg, the project team only conducted ART initiations; HIV testing was conducted by another NGO team as part of a Global Fund–supported service. In the 2 Western Cape sites, the project team ensured that ART medications were available in all instances, but UTT services (testing and treatment initiation) were delivered by the DCS clinic team as part of routine care. In all sites, incarcerated respondents’ and DCS stakeholders’ accounts demonstrated the feasibility of delivering UTT within these particular South African correctional facilities, with interview data and investigator experience pointing to 2 major contributing factors as described below.

National UTT Policy and Study Resourcing

Prior to project start on September 1, 2016, the South African National Department of Health (NDOH) introduced a national UTT policy, which provided the criteria for starting all patients with HIV on lifelong ART. Since DOH guidelines for the management of HIV apply to correctional centers in South Africa, this change in national policy was widely understood to apply to correctional settings. Interviews highlighted how knowledge of this policy among both frontline and administrative DCS staff facilitated the establishment of UTT in correctional facilities. Despite some early fears among some to the contrary, DCS staff at all 3 sites additionally described how the study helped them absorb the backlog of counseling and paperwork associated with initiating many now-eligible individuals with HIV into HIV treatment.

UTT, when it was initiated by the Department of Health, it became something very crucial for us [in corrections] But now, there’s no backlog anymore [. . . the study made] it easier for us to get those inmates on
treatment as soon as they were being diagnosed. [Nursing sister, South Africa]

Study resources, which included personnel to help with HIV testing and counseling and importantly antiretroviral medications for newly initiated individuals at the study sites, as well as the development of tailored protocols, helped DCS health workers make the transition to UTT in the 2 study sites with relatively few disruptions to existing duties.

**Correctional Health System and Established Partnerships**

Respondents’ descriptions of the health services in the 3 South African sites, combined with the investigators’ experience, point to robust systems for the delivery of health care in the 2 South African correctional facilities. In particular, transparent processes for accessing HIV testing and treatment provided a platform from which UTT procedures could evolve. As a frontline provider noted:

*Prior to this project being implemented there was a certain working procedure towards HIV/AIDS and TB treatment, care and support [developed with our partners]. I think over the years we have, the wheels were well oiled. [Nursing sister, South Africa]*

Reflecting the comparatively strong internal capacity of the South African correctional system, several DCS staff clearly positioned themselves as champions of UTT, serving as informal and formal conduits of information and advocates for necessary changes or adaptations to other DCS staff. Preexisting relationships between the (separate) NGO implementing partners at the 2 study sites and DCS officials also provided an important basis for communication and problem solving, helping overcome early challenges and misunderstandings.

**Sustainability of UTT in South African Correctional Facilities**

The prospects for sustained UTT in the 3 South African correctional facilities appeared strong, although barriers remained. Reflecting both deductive and inductive themes, we identified barriers and facilitators relating to 4 major sustainability domains as outlined below.

**Policy Backdrop and Political Support:** As described above, South Africa’s national UTT policy was key to promoting acceptance of the TasP project. But as noted in the quote below, the policy also ensured that DCS staff were interested in facilitating and adopting UTT in the long term.

*Like I said with regards to UTT when it was initiated by the Department of Health it became something very crucial for us [to implement]. [DCS officer]*

This understanding that the impetus for UTT came from central government, rather than from the study itself, also contributed to buy-in by mid- and senior-level DCS officials in relation to the longer-term adoption of UTT. As one noted:

*The study helped us a lot [to] see how this is actually feasible. But even if [study funding] stops I think we must continue with [the] way we are working now. [Correctional HIV AIDS Coordinator, South Africa]*

However, some barriers to UTT sustainability were noted. Several health care providers described current DCS policies that do not empower nurse prescribing and ART initiation as a potential impediment to sustained and effective UTT, since ongoing shortages of medical doctors in the correctional system could lead to bottlenecks in treatment initiation.

**Funding and Resourcing:** Notwithstanding the relatively well-established systems in the South African correctional health services, a common theme was the “brittleness” of these services. Data revealed that corrections health services operated on a skeleton staff who had little time to deal with anything beyond routine check-ups. Frontline health workers at all 3 sites described chronic (reportedly system-wide) staff shortages that necessitated a daily juggling act as they tried to deliver services for both acute and chronic conditions across a large population with multiple, often complex physical and mental health care needs.

*Nurses are under staffed again and over worked because it’s a lot of [health] complaints. … I think that [shortage] can have a very negative effect on the health care service delivery. [Correctional HIV AIDS Coordinator, South Africa]*

The implications for understaffing for UTT were many, but one standout consequence emerging from interview data was the inability of DCS to deliver on critically important psychosocial counseling during initial HIV testing or as part of ongoing HIV treatment. DCS-employed HIV counselors or other psychosocial supporters were limited. Yet their services were described independently by both incarcerated individuals and providers as essential to incarcerated individuals’
uptake of and sustained engagement in HIV care and treatment. Favorable comparisons were made by respondents at the various sites to the counseling made available during the study, when TasP staff with dedicated counseling roles bolstered health worker numbers:

“It’s not like taking chocolates and everything, you know. You’ve got to be, ja [hand gesture]. And then we received such perfect, perfect counseling with [the TasP team]. And it makes it so easy for us, you know, to, to take our pills each and every day. Because we know the purpose, we know why we are taking these pills. Ja man, prison! [Male inmate, Cape Town]

Organizational Capacity and Organizational Culture: The comparatively well-established health services within DCS were facilitators of scaled-up and sustained UTT across the 3 South African correctional sites. Nonetheless, 2 corrections-specific organizational barriers were identified. The first, related to health information challenges. These included still weak internal health information systems and resulting communication breakdowns during transfers of patients between corrections facilities. Inconsistent communication between security and health personnel and nonharmonized health information systems between facilities were described as often resulting in missing or delayed medical records that left patients on ART without access to medication. At a higher level, lack of harmonization between DCS and DOH indicators for HIV care and treatment were identified as likely inhibiting long-term monitoring.

The second aspect of organizational capacity related to the weak health literacy of security personnel and its impact on service organization and access. As described by one security officer:

“We did not receive training on the college to work with [NGO] personnel or with providing of treatment to a person. We did not get that kind of training at Corrections College. Uhm, that is why some of the [staff] felt that: “it’s not my job description. Where do these people come from now?” [Security officer, South Africa]

Coordination between DCS security and health personnel were described as challenging in some situations, potentially limiting incarcerated peoples’ access to care. Efforts to ensure leadership and communication between different categories of corrections staff were described as an area requiring improvement to facilitate sustained and effective UTT over time.

Population Dynamics and Service Demand: A final theme relating to UTT sustainability in the South Africa corrections health system relates to population dynamics of each facility and their effect on UTT service demand. Incarcerated individuals, health workers, and DCS officials all described the correctional facilities as high-stress environments in which peer relations among the incarcerated were fraught and emotional violence was common.

Some of our clients … suffer a lot with, for instance, depression. Because you know the prison system, the correctional system, is not a place where someone wants to be … their freedom is being taken from them, their families, their loved ones, their children can’t see them every day, … And now you come and stay with total strangers, in a strange place … the medical or the psychological effect that it has on the inmates is something that we can’t deny. [Nursing sister, Gauteng]

Many incarcerated respondents and health providers described how, in this stressful and isolating environment, self and perceived stigma around HIV status was a barrier to accessing or remaining engaged in HIV treatment. HIV-related stigma was a prominent concern among the majority of the individuals with HIV who were interviewed, and it was confirmed by health care providers. Self-stigma, typically expressed as a reluctance to reveal or disclose to anyone else, was also frequently described, exacerbated by poor HIV treatment literacy, lack of personal coping mechanisms, and myths and misconceptions around the effects and implications of starting ART.

It is a challenge, it is a challenge, because at first when [stuttering] you are told that you’re actually now having to live with the fact you are HIV, you know, your mind shuts [down] immediately. You know, you [stuttering] get confused and you think of [stuttering], the environment now that you are in. I think it would be much better if I was outside, but now I’m in prison, you know, where we are ten inmates in a small [single] cell, and these people will see that I’m taking pills. [Male inmate, Gauteng]

Individually and in combination, these factors were consistently linked by different types of respondents to poor uptake of existing (pre-UTT) HIV testing and to poor adherence to ART.

We are not on 100% compliance with regards to ARV treatment even now. Not because it’s not available … Some of them said [they don’t want to start treatment] because of their family. Their culture says they mustn’t. Others say they just don’t want it. But if you go and dig a
Some women who had chosen not to start ART, raised concerns regarding the “double sentence”—physical imprisonment and the “sentence” of HIV. These individuals explained that they did not want to have to worry about HIV or initiate treatment until they were released. Against this backdrop, the weak presence of HIV-specific or more generalized psychosocial counseling or support services in the South African correctional health services was described as a substantial limitation to sustained and effective UTT.

### DISCUSSION

Recent expansion of HIV testing and treatment in heavily HIV-burdened countries has been framed as a global imperative. However, the key populations designated by Joint United Nations Program on HIV and AIDS as being at greatest risk for HIV, including incarcerated peoples, are being left behind in the global response, often going without proper access to HIV treatment and other health services. While large-scale clinical trials have demonstrated the effectiveness of new biomedical technologies for tackling HIV, a number of unanswered questions remain, particularly with regard to how we implement these technologies for certain populations and in various subnational contexts. Indeed, information about or evidence from interventions to help address HIV care and treatment among incarcerated populations in sub-Saharan Africa and elsewhere, remains limited.

To our knowledge, and notwithstanding the overwhelming burden of HIV in sub-Saharan Africa, this study is the first to qualitatively explore the feasibility and sustainability of UTT in any African correctional system. A key contribution of this study is its consideration of both the interventional and contextual factors likely to affect UTT, synthesized through comparison of the influence of the different political, correctional, and health service delivery systems in Zambia and South Africa, respectively. It is also one of still very few articles from the region describing any form of health intervention in the correctional context. Given the dearth of literature, we focus the Discussion on 3 major, cross-cutting lessons synthesized from the findings presented above, which are relevant to the broader study questions but also important flags for policy makers and programmers who may be considering UTT roll-out in correctional systems elsewhere.

### National Policies and Associated Resourcing Are Critical

Both the Zambian and South African experiences in this study demonstrated the critical influence of national HIV treatment policies on the feasibility and sustainability of UTT in correctional health systems, albeit in different ways. In South Africa, the national policy was interpreted as inclusive of incarcerated populations, influencing the response of DCS officials and facility-level providers who felt compelled to implement UTT in correctional facilities. In Zambia, the national UTT policy ensured general support for the project. However, the lack of clarity at a high level about whether the policy—a form of clinical guidance—guaranteed the necessary resources had implications not only for the capacity of ZCS to implement it, but also for correctional officials’ sense of urgency or compulsion to comply with it. These experiences highlight the importance of understanding both the political and resourcing context of policy change, and they mirror concerns elsewhere about the lack of specificity in nominally universal health policies, which can falter in operationalization due to differing interpretations of scope, intent, or measurement.

Although less apparent in our primary data, author experience and a critical body of literature on correctional health in sub-Saharan Africa point to the key role that reform of the justice system (e.g., bail alternatives to limit over-incarceration) could play in making HIV prevention and treatment more effective and affordable. Such reforms are central to tackling upstream drivers of incarcerated people’s lack of access to care (e.g., by reducing overcrowding through parole reform) and constitute a fundamental and complementary set of strategies in the quest for sustained deliver of UTT.

### Correctional Health Delivery Systems and Organizational Capacity

The influence that the capacity of the correctional health system had on the feasibility and sustainability of UTT in correctional facilities is an intuitive but central finding. In Zambia, some recent progress in establishing internal health systems was noted, but it was insufficient to independently support and sustain UTT. Chronic shortages...
The narratives of incarcerated individuals pointed to intersectionalities between mental health, stigma, and HIV.

An urgent need exists to boost human resource capacity and ARV stocks to manage testing and treatment among incarcerated people.

HIV Stigma and Service Demand
Reflecting the importance of health system software on the feasibility and sustainability of UTT in corrections, we found HIV stigma and associated service demand within correctional facilities to be a key cross-cutting theme. In the Zambian female section and in South African facilities, respondents described HIV-related stigma as an ongoing barrier to UTT uptake. Accounts referenced recent and past trauma that interacted with feelings of fear and isolation, leaving many individuals unwilling to start treatment even where they knew that testing and treatment were accessible.

Our findings align with previous research among incarcerated males in South Africa that has traced the psychosocial determinants of risk behavior and in which respondents’ narratives of fear and isolation segued into descriptions of (emotional and psychological) barriers to health care access. Various (noncorrectional) studies have also traced the influence of psychological, social, and cultural factors on HIV care seeking and willingness to enroll in ART in other sub-Saharan African countries. This qualitative study, with its focus on correctional health systems, was not specifically designed to assess mental health, although a mental health assessment of enrollees with HIV in the clinical study was conducted and will be reported elsewhere. Nonetheless, the narratives of incarcerated individuals from both countries pointed to the likely intersectionalities between incarcerated people’s mental health, stigma, and HIV. From a service and treatment efficacy perspective, it seems reasonable to assume that poor mental health among incarcerated people affects their agency to seek and remain engaged in care. Future work should more explicitly explore these links and their implications for the profile and the skills mix of health care services and providers in correctional settings.

Countercurrently, interviews with incarcerated males in the Zambian facility in this study described a strong culture of peer support for HIV testing and treatment, describing it as an important factor in their willingness to seek care, even where physical access was limited. As recent work has suggested that key populations in Zambia continue to face barriers to service utilization in the mainstream health system, more work is needed to explore and trace the factors that may have enabled and strengthened this unusual and encouraging situation in the correctional facility.

Limitations
This study analyzed data relating to the implementation of UTT in 1 Zambian facility and several South African facilities. While we endeavored to garner views representative of different levels of the health system—client, frontline health worker, managers, and policy makers—in order to improve analytic generalizability, we acknowledge that both Zambia and South Africa have multiple correctional facilities that differ in size, security level, infrastructure, and administrative dynamics.

Investment in the longer-term health system management and planning capacity building for frontline officers and mid- and upper-level ZCS officials nonetheless remains an important priority.

South Africa’s correctional health services showed capability to take on and adopt UTT. However, mirroring the experience of UTT implementation in mainstream South African health services, the study also revealed a number of issues regarding scalability and sustainability. Reflecting concerns in other recent literature on South Africa’s correctional system, our findings demonstrated an urgent need to boost human resource capacity and ARV stocks to manage the surge of testing and treatment initiation created by the backlog of treatment-eligible incarcerated people. The study also revealed how current national clinical guidelines that require a medical doctor to initiate ART are already contributing to treatment bottlenecks in correctional facilities where doctors are few, a situation likely to be exacerbated by UTT. Limited provision of essential pre- and posttest counseling and psychosocial support for clients with HIV were also linked to gaps in skills mix of the correctional health workforce. Finally—similar to, although less acute than in Zambia—findings highlighted the need to strengthen correctional health information systems to enable better linkage and tracking between correctional facilities and across the corrections and mainstream health systems.
from the study sites. Zambia and South Africa do not prevent noncitizens from accessing health services in corrections, so we did not interview participants to reflect on their ability to access UTT. However, we did not interview any confirmed noncitizen incarcerated persons in either country, and we acknowledge that their experience of correctional services and ability to access UTT within such settings may well be different from that of citizens. Because the larger study in which this qualitative work was nested focused on implementation of UTT, we also did not focus explicitly on other service or upstream gaps (e.g., mental health and/or criminal justice reform) future work in corrections should seek to include such gaps in these and similar settings. Further work to identify the synergies of such services or reform efforts for ensuring (among many outcomes) long-term sustainability of HIV-specific services will be important.

## CONCLUSION

Despite different institutional profiles, introduction of UTT was feasible in correctional health systems in both South Africa and Zambia. A supportive policy backdrop was important to UTT establishment in both countries. However, the prospects for long-term sustainability of UTT differed. South Africa’s correctional facilities had existing systems to deliver and monitor chronic HIV care and treatment, forming a scaffolding for sustained UTT despite some human resources shortages and poorly integrated health information systems. Notwithstanding recent improvements, Zambia’s correctional health system demonstrated insufficient material and technical capacity to independently deliver quality UTT. In the correctional facilities of both countries, population dynamics and their impact on HIV-related stigma were important factors in UTT service uptake.

Acknowledgments: We acknowledge the participation and contributions of incarcerated persons, correctional officials, and national policy makers in Zambia and South Africa, which made the TasP project and this study possible.

Funding: This research was funded through a grant from the DFID and SIDA-funded Evidence for HIV Prevention in Southern Africa (EHPSA) program: Grant Number: MM/09/EHPSA/AURUM/051.500113

Competing Interests: None declared.

## REFERENCES


Peer Reviewed

Received: January 30, 2019; Accepted: April 9, 2019


© Topp 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-19-00051
Alternative Ready-To-Use Therapeutic Food Yields Less Recovery Than the Standard for Treating Acute Malnutrition in Children From Ghana

Kristin Kohlmann, Meghan Callaghan-Gillespie, Julia M. Gauglitz, Matilda Steiner-Asiedu, Kwesi Saalia, Carly Edwards, Mark J. Manary

In Ghana, an alternative ready-to-use food (RUTF) formulation that met all specifications was not as good as standard RUTF in affecting recovery from acute malnutrition among children aged 6 to 59 months.

ABSTRACT

Background: Only 20% of children with severe acute malnutrition (SAM) have access to ready-to-use therapeutic food (RUTF), and RUTF cost limits its accessibility.

Methods: This randomized, double-blind controlled study involved a clinical equivalence trial comparing the effectiveness of an alternative RUTF with standard RUTF in the home-based treatment of uncomplicated SAM and moderate malnutrition in Ghanaian children aged 6 to 59 months. The primary outcome was recovery, equivalence was defined as being within 5 percentage points of the control group, and an intention-to-treat analysis was used. Alternative RUTF was composed of whey protein, soybeans, peanuts, sorghum, milk, sugar, and vegetable oil. Standard RUTF included peanuts, milk, sugar, and vegetable oil. The cost of alternative RUTF ingredients was 14% less than standard RUTF. Untargeted metabolomics was used to characterize the bioactive metabolites in the RUTFs.

Results: Of the 1,270 children treated for SAM or moderate malnutrition, 554 of 628 (88%) receiving alternative RUTF recovered (95% confidence interval [CI]=85% to 90%) and 516 of 642 (80%) receiving standard RUTF recovered (95% CI=77% to 83%). The difference in recovery was 7.7% (95% CI=3.7% to 11.7%). Among the 401 children with SAM, the recovery rate was 130 of 199 (65%) with alternative RUTF and 156 of 202 (77%) with standard RUTF (P=.01). The default rate in SAM was 60 of 199 (30%) for alternative RUTF and 41 of 202 (20%) for standard RUTF (P=.04). Children enrolled with SAM who received alternative RUTF had less daily weight gain than those fed standard RUTF (2.4±2.4 g/kg vs. 2.9±2.6 g/kg, respectively; P<.05). Among children with moderate wasting, recovery rates were lower for alternative RUTF, 386 of 443 (87%), than standard RUTF, 397 of 426 (93%) (P=.003). More isoflavone metabolites were found in alternative RUTF than in the standard.

Conclusion: The lower-cost alternative RUTF was less effective than standard RUTF in the treatment of severe and moderate malnutrition in Ghana.

INTRODUCTION

In sub-Saharan Africa, 17 million children under 5 are wasted, which is defined as having a weight-for-length z score (WLZ) < 2 standard deviations (SD) below the mean World Health Organization (WHO) Child Growth Standards. Wasting leaves these children with an increased risk of illness and death. A large fraction of wasting occurs in children aged 6 to 24 months, a dynamic period of physical and neurological development. The majority of wasted children do not live in communities beset with emergencies, but rather come from the poorest segments of all countries. In general, these countries do not have the resources from donated or endogenous sources to sponsor widespread feeding and education programs to combat wasting. Effective and cost-efficient solutions to reduce wasting outside of acute emergencies will be necessary to achieve the Sustainable Development Goals.

Among children under 5 years of age in Ghana, the prevalence of wasting is about 5%. Treatment for severe wasting, which is defined as WLZ ≤ −3 SD below the mean, is available in the northernmost regions of Ghana, where the density of SAM is greatest but where
only 17% of the population reside. Treatment of moderate wasting, defined by WLZ > −3 and ≤−2 and known as moderate acute malnutrition (MAM), is almost entirely unavailable in Ghana.

Home-based therapy with ready-to-use therapeutic food (RUTF) for children with SAM has revolutionized the management of wasted children, offering a superior alternative to inpatient treatment.4,5 Unfortunately, RUTF reaches only about 15% of the children worldwide who need it. Despite being highly cost-effective, SAM treatment is expensive in absolute terms, with a cost of US $150 to $200 per child, and in Ghana, one limited study estimated the cost of treating SAM to be $805 per child.6,7 Worldwide, standard RUTF (S-RUTF) is an expensive component of treatment, costing $47 to $61 per child treated.8 S-RUTF is composed of 25% skimmed milk powder and 27% peanut paste, a vegetable oil rich in omega-3 polyunsaturated fatty acids such as canola and sugar.

In 2013, our team initiated a multinational alternative RUTF (A-RUTF) formulation project with the aim to reduce the cost of RUTF, and in doing so, enable the existing resource envelope for SAM to be used to treat more children. The work began with a comprehensive literature and nutrient database analysis and subsequent development of a food formulation linear programming (LP) tool.9 The LP tool is a conventional computer database program that lists all potential ingredients, nutritional compositions, prices, and country-specific availability. The tool has default nutrient constraints that align the formulations with the international RUTF nutrient specifications and food safety guidelines.10 The tool also allows for ingredient constraints, which supports organoleptic optimization.5,11 It has been successfully used by our investigative team to create country-specific locally produced A-RUTF formulations for Ghana, Ethiopia, Pakistan, and India that were proven to be feasible, acceptable, and without adverse side effects in formal acceptability trials.12 However, the relative effectiveness of an A-RUTF to S-RUTF has yet to be shown.

This article describes the operation and results from a randomized, double-blind controlled clinical trial testing the hypothesis that a locally produced A-RUTF was equivalent to S-RUTF for the treatment of uncomplicated SAM and MAM, in the Brong Ahafo region of Ghana.

**METHODS**

**Subjects and Setting**

Eligible children were between 6 and 59 months of age and experiencing acute malnutrition. SAM was defined as WLZ ≤ −3, or having a mid-upper arm circumference (MUAC) of <11.5 cm or bipedal edema. MAM was defined as not having SAM and having WLZ ≤ −2 or MUAC of <12.5 cm. In addition to meeting the anthropometric criteria, children were required to consume 30 g of RUTF in a supervised setting to be eligible for enrollment. Children were excluded if they were involved in another research trial or feeding program, had a chronic debilitating illness (e.g., cerebral palsy), or had a history of peanut or milk allergy.

Informed consent was obtained from the primary caregiver of the participant and documented by the caregiver’s signature or thumbprint. The study received ethical approval from the Washington University in St. Louis Institutional Review Board, the Noguchi Memorial Institute for Medical Research Institutional Review Board, and the Ghana Health Service.

Study participants were recruited at 29 clinics throughout 5 districts in the Brong Ahafo region of Ghana. The Brong Ahafo region is the second largest region in Ghana and has the sixth largest population at 2.3 million.13 In 2011, the under-5 mortality rate in the region was 108 deaths per 1,000 live births, 32% higher than the national under-5 mortality rate.14 Although wasting rates in the latest Demographic and Health Survey showed a national decline, regional trends indicated that rates in the Brong Ahafo region had increased.15 In addition, 16% of all households in this region are considered food insecure.16

**Study Design**

This randomized, double-blind controlled study was based on a clinical equivalence trial of treating acute malnutrition with 1 of 2 therapeutic foods, A-RUTF or S-RUTF. The primary outcome was recovery, defined as having achieved either WLZ > −2 or MUAC >12.4 cm at any point during the treatment. Equivalence was chosen as being within 5 percentage points of the control group. Secondary outcomes were rates of weight and MUAC gain, number of visits before recovery, cost of RUTF per child recovered, and adverse events. The sample size was estimated to be 1,262 children, which gave the comparison sufficient power to detect a 5% difference in recovery, assuming the control group achieved recovery rate of 85% using an equivalence design. The assumption that recovery would be 85% overall for the treatment of MAM and SAM was based on our trials in Malawi.8 The trial was publicly registered as ISRCTN14788669.
Participation and Data Collection

All participants were randomized to receive either A-RUTF or S-RUTF via a closed envelope technique. Allocation of the food intervention was conducted by a nurse who had the participant’s caregiver draw an opaque envelope containing 1 of 4 colors. Each color corresponded to a type of RUTF. Both the research team and study participants were blinded to color assignments.

Management of MAM and SAM followed an optimized protocol that incorporated many elements from the community management of acute malnutrition (CMAM), which is described in Table 1. Notable deviations from CMAM were (1) visits were fortnightly instead of weekly, (2) the ration of RUTF for SAM was reduced as the child gained weight, (3) MAM children were given supplementary food in addition to counseling, and (4) exit criteria for the study were achievement of MUAC >12.4 cm on a single occasion or completion of 12 weeks of feeding, instead of requiring 3 occasions with MUAC >12.4 cm.

The children had MUAC, weight, and length measured upon enrollment. MUAC was measured on the left arm with a standard insertion tape to the nearest 0.1 cm (TALC, Herts, UK); weight was measured to the nearest 5 g using an electronic scale (Seca 334, Hamburg Germany, calibrated weekly); and recumbent length was measured in triplicate to the nearest 0.2 cm, using a rigid length board (Seca 417 length board, Hamburg, Germany). The staff received standardized training every 8 weeks in the measurement of edema and anthropometry by a senior clinician, and 10% of the field measurements were rechecked in the field for quality purposes. During the initial visit, demographic and health information were recorded, and a 2-week supply of their assigned RUTF was dispensed. The dosage of RUTF provided a daily intake of about 150 kcal/kg for SAM participants and about 75 kcal/kg for MAM participants. The daily SAM ration provided about 100% of the child’s needs for growth and maintenance and was typically about 200 g. The daily MAM ration provided about 60% of the child’s needs for growth and maintenance and was typically about 100 g. Caregivers and study participants were asked to return every 2 weeks for follow-up. At follow-up, caregivers reported on the child’s clinical symptoms, anthropometric measurements were taken, and additional RUTF was distributed for those that remained wasted. The dosage of RUTF distributed at each follow-up visit was determined by the child’s current weight. As SAM participants began to recover and reached a MUAC ≥11.5 cm, they were transitioned to the MAM dosage of 75 kcal/kg/day of their assigned RUTF. No additional food rations were given when subjects reached an outcome, nor were the children asked to return for follow-up at regular intervals.

The study was implemented by trained nurses working for Project Peanut Butter, a registered NGO in Ghana. A research associate from Washington University resided in Ghana for the

| TABLE 1. Comparison of Project Peanut Butter and Ghana Health Service Malnutrition Management Protocols in Brong Ahafo, Ghana |
|---------------------------------|-------------------------------------------------|-------------------------------------------------|
| **Project Peanut Butter Protocol** | **Ghana Health Service CMAM Protocol** |
| **SAM enrollment criteria** | MUAC <11.5 cm or WLZ < −3 SD Bilateral pitting edema | MUAC <11.5 cm or WLZ below −3 SD Bilateral pitting edema |
| **MAM treatment** | Enrolled and treated with RUTF | Increased nutrition counseling during CWC |
| **MUAC ≥11.5 cm, <12.5 cm** | 150 kcal/kg/day for SAM 75 kcal/kg/day for MAM | 200 kcal/kg/day for SAM |
| **WLZ between −2 and −3 SD** | Biweekly | Weekly |
| **RUTF dosage** | Graduation criteria | Discharge criteria |
| **Follow-up** | MUAC >12.4 cm, or WFL > −2 SD for 1 visit (2 weeks) | 3 consecutive visits missed (6 weeks) |
| Maximum duration of treatment | 12 weeks | 16 weeks |

Abbreviations: CMAM, community management of acute malnutrition; CWC, Child Welfare Clinics; MAM, moderate acute malnutrition; MUAC, mid-upper arm circumference; RUTF, ready-to-use food; SAM, severe acute malnutrition; WLZ, weight-for-length z score.
Both RUTF formulations used in the study met nutritional and microbiological requirements and underwent safety testing.

Study Foods
Both RUTFs were produced at Project Peanut Butter in Kumasi, Ghana, a certified local supplier. Both RUTF formulations met the nutritional specifications and microbiological requirements for RUTF set forth by United Nations agencies in 2007 and underwent safety testing for aflatoxin and microbial contamination at Eurofins Scientific Inc. (Des Moines, Indiana, USA). The S-RUTF contained peanut paste, sugar, nonfat dried milk, vegetable oil, a premix containing concentrated minerals and vitamins, and an emulsifier. The A-RUTF replaced about half the amount of peanut with locally available soybean and sorghum flour, and the 50% of protein from dairy per United Nations specification came from a combination of whey protein concentrate 34 and nonfat dried milk. A-RUTF also included canola oil, sugar, a vitamin and mineral premix, as well as less nonnutritive emulsifier (Table 2).

Whenever study food was given to children, the nurses counseled the caregivers to feed the RUTF in whatever manner the child would readily accept it, which was most often sucking the food out of the flexible package from a small tear. Caregivers were also counseled to feed the RUTF strictly to the maldnourished child and not to share or sell the RUTF.

The ingredient cost of A-RUTF was US$1.90/kg compared with $2.20/kg for the S-RUTF, a 14% cost reduction in ingredients. This reduction was largely achieved by substituting the less expensive sorghum and soy for peanut. Ingredient prices were estimated using the LP tool, which employed a modeling method that determined the median commodity prices in 2012 in Ghana from a comprehensive variety of sources, including accounting for transportation and taxes. The price variation seen in the subsequent 5 years was then added to the model to estimate “typical” prices for the ingredients.

Protein quality was calculated to better characterize A-RUTF and S-RUTF. The Digestible Indispensable Amino Acid Score (DIAAS) method with the reference population being healthy children aged 1–3 years was used to calculate protein quality. In addition, the DIAAS was recalculated using malnourished children in a phase of rapid catch-up growth as a reference population.

Coverage Survey
To determine coverage of MAM and SAM children receiving RUTF feeding (i.e., the proportion of children with acute malnutrition who were accessing services), we used the simplified lot quality assurance sampling evaluation of access and coverage (SLEAC) method. The coverage survey was conducted as a routine measure of program effectiveness, which allowed us to understand if the research feeding achieved similar coverage as operational programs in sub-Saharan Africa.

Metabolomics Analysis
To characterize the nonnutritive components of the RUTFs, which might contribute to the clinical effect, untargeted metabolomics analyses were conducted. A-RUTF and S-RUTF were extracted to a final concentration of 1 µg/µL in 50% methanol and 95% ethanol for untargeted metabolite analysis. Data were acquired for each sample in triplicate using an ultra-high performance liquid chromatography–tandem mass spectrometry system (UltiMate 3000 UHPLC system [Thermo Scientific, Waltham, MA, USA] coupled to a Maxis Q-TOF mass spectrometer [Bruker Daltonics, Bremen, Germany]), using electrospray ionization in positive mode and a reverse phase C18 column (Kinetex, 100 × 2.1 mm, 1.7-µm particle size, 100-Å pore size; Phenomenex, Torrance, CA, USA). Raw data files were converted to mzXML format using Bruker DataAnalysis software after lock mass correction (m/z=622.0290; Hexakis [SynQuest Laboratories, Alachua, FL, USA]) and analyzed with molecular networking and library spectral matching using the web-based platform GNPS (https://gnps.ucsd.edu). The analysis is available at https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=a474e2ed686f43d7b2946a53225495c2.

Data Analysis
Data were double entered into a Microsoft Access database and discrepant values corrected by reviewing the original data collection cards. For children older than 24 months, height was estimated from the measured length by subtracting 1.5 cm from the length. Z scores were calculated using the WHO Anthroplus version 1.0.4 (WHO, Geneva), based on the 2006 WHO Child Growth
Rates of weight gain were calculated for the first 4 weeks of treatment by dividing weight gain in grams by the enrollment weight in kilograms and the days of treatment between measurements. Mean daily MUAC gain was also calculated for the first 4 weeks of treatment by dividing MUAC gain in millimeters by days of treatment between measurements.

Data were analyzed by using SPSS Statistics software (version 25.0; IBM Corp., Armonk, NY, USA). Summary statistics were calculated for the participants as mean ± SD for continuous parameters and n (%) for categorical parameters. Analyses were done by intention to treat (ITT) for which defaulters were considered to be failures in accordance with the Sphere Standards. In accordance with the trial designation as an equivalence trial, the 95% confidence intervals (CIs) around the recovery rates were calculated to determine if there was overlap between the groups and the difference was compared to determine if it exceeded the threshold of 5 percentage points.

Subgroup analyses were performed on children with SAM and MAM. For secondary and subgroup outcomes treatment groups were compared using the Student’s t test or Mann-Whitney U test for continuous variables and Fisher’s exact test for categorical measures.

## RESULTS

A total of 1,270 children were enrolled in the study from February 2017 to February 2018 (Figure 1). Of these, 401 were diagnosed with SAM and were assigned to receive either A-RUTF (n=199) or S-RUTF (n=202); 869 children were diagnosed with MAM and were assigned to receive either A-RUTF (n=443) or S-RUTF (n=426). The baseline

### TABLE 2. Ingredient and Nutrient Composition of Study Foods

<table>
<thead>
<tr>
<th>Ingredient/Nutrient</th>
<th>Alternative-RUTF</th>
<th>Standard-RUTF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ingredient</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cereal/grain, sorghum, g/100 g</td>
<td>9.00</td>
<td>—</td>
</tr>
<tr>
<td>Legume, g/100 g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Groundnut</td>
<td>14.00</td>
<td>27.00</td>
</tr>
<tr>
<td>Soybean</td>
<td>2.00</td>
<td>—</td>
</tr>
<tr>
<td>Milk, g/100 g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dry, nonfat, regular, without added vitamin A and vitamin D</td>
<td>5.00</td>
<td>25.00</td>
</tr>
<tr>
<td>Whey protein concen</td>
<td>20.18</td>
<td>—</td>
</tr>
<tr>
<td>Oil, g/100 g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canola</td>
<td>20.50</td>
<td>—</td>
</tr>
<tr>
<td>Palm</td>
<td>—</td>
<td>15.48</td>
</tr>
<tr>
<td>Soybean</td>
<td>—</td>
<td>2.92</td>
</tr>
<tr>
<td>Sugar, g/100 g</td>
<td>25.00</td>
<td>24.64</td>
</tr>
<tr>
<td>Micronutrient and vitamin premix, g/100 g</td>
<td>2.92</td>
<td>2.96</td>
</tr>
<tr>
<td>Emulsifier, g/100 g</td>
<td>1.40</td>
<td>2.00</td>
</tr>
<tr>
<td>Nutrient</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy, kcal/100 g</td>
<td>560</td>
<td>559</td>
</tr>
<tr>
<td>Protein, g/100 g</td>
<td>14.5</td>
<td>15.8</td>
</tr>
<tr>
<td>Lipids, g/100 g</td>
<td>29.2</td>
<td>33.0</td>
</tr>
<tr>
<td>n-6 fatty acids, g/100 g</td>
<td>6.3</td>
<td>5.7</td>
</tr>
<tr>
<td>n-3 fatty acids, g/100 g</td>
<td>1.9</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Abbreviation: RUTF, ready-to-use therapeutic food.
* Both foods were a soft, brown, homogeneous paste with small granules perceptible to the tongue. They were packaged in identical, unlabeled metalized polyethylene terephthalate sachets with the only marking being a colored dot to indicate the type of RUTF.
characteristics for each study group were similar (Table 3). For the children with SAM, 157 were designated by both MUAC and weight-for-length, 178 only by MUAC, 63 only by weight-for-length, and 3 only by having edema. For the children with MAM, 320 were designated by both MUAC and weight-for-length, 378 only by MUAC, and 171 only by weight-for-length.

Among SAM and MAM children receiving A-RUTF, 516 of 642 recovered (80.4%, 95% CI= 77.1% to 83.3%) (Table 4). Among children receiving S-RUTF, 553 of 628 recovered (88.1%, 95% CI= 85.3% to 90.4%). The difference in recovery rates was 7.7 percentage points (95% CI=3.7 to 11.7 percentage points).

The protein content in 100 g of S-RUTF was 1.25 g higher than in 100 g of A-RUTF, although both foods were within current international agency specifications and 50% of protein was from a dairy source (Table 2). S-RUTF also had a lipid content that was 2.8 g more than A-RUTF per 100 g of RUTF. Protein quality as determined by DIAAS using healthy children as a reference population was 107 for A-RUTF and 109 for S-RUTF. Protein quality using malnourished children as a reference population was 85 for A-RUTF and 84 for S-RUTF.

Among SAM children, recovery was seen in 130 of 199 (65.3%) for those receiving A-RUTF and was 156 of 202 (77.2%; P=.01) for those receiving S-RUTF (Table 4, Figure 2); defaults were not considered to be recovered in our ITT analysis. Children with SAM receiving A-RUTF had lower rates of gain for both weight and MUAC (Figure 2). Children with MAM who received A-RUTF were less likely to recover, according to an ITT analysis (Figure 3). For children with MAM, rates of weight and MUAC gain were similar between groups (Table 4). Mortality rates were low, with only 5 (0.4%) MAM or SAM children dying.

A total of 184 (14.5%) of children did not complete the study and were classified as defaults (Table 4). The relative risk for defaulting if the child was enrolled with SAM compared with those
with MAM was 3.39 (CI 95%: 2.53 to 4.53). Of the children who defaulted, 116 (63.0%) received A-RUTF compared with 68 (37.0%, \(P < .001\)) who received S-RUTF. Considering the SAM children who defaulted, 68 of 101 (67.3%) did so before the 4-week follow-up, and 7 of 101 (6.9%) did so after the 8-week follow-up. Only 26 of 101 (25.7%) attained an MUAC > 11.4 cm, indicative of improvement from SAM to MAM.

Coverage surveys were conducted throughout catchment areas in February 2018. During the survey, the data collection teams assessed a total of 560 children. Among these children, 116 (20.0%) had SAM and 52 (9.3%) had MAM. The coverage of SAM children was 7 of 11 (63.6%) and MAM children was 18 of 52 (34.6%). Mothers were asked if they were aware there was a treatment program in their community, and 8 of 11 (72.7%) of mothers with SAM children and 28 of 52 (53.8%) of mothers with MAM children responded positively.

The cost of A-RUTF used per MAM child recovered was US$7.07, while for S-RUTF the cost was US$8.20 (16% higher). The cost of A-RUTF per SAM child recovered was US$28.72, while for S-RUTF this was US$28.48, a similar amount.

Untargeted metabolomics of A-RUTF and S-RUTF showed that among the 26 unique metabolites found in A-RUTF, 5 were isoflavones, consistent with the addition of soy products in A-RUTF; while S-RUTF had only 9 unique metabolites, and most were likely to be minor components of the food emulsifier or peanuts (Table 5). No xenobiotics were found in the A-RUTF that were also not present in the S-RUTF.

**DISCUSSION**

An acute malnutrition treatment program was successfully instituted at 29 rural sites in Brong Ahafo, where the prevalence of acute malnutrition was high and a coverage estimate met those typically reported as well as international standards.\(^2^3\),\(^2^4\) Unexpectedly, A-RUTF was not equivalent to S-RUTF in the treatment of SAM or MAM in Ghana in this randomized, double-blind, clinical, controlled trial compared with an ITT analysis. The primary compositional differences were that sorghum and soy were used in A-RUTF in place of some of the peanut paste in S-RUTF and a large portion of the dried skim milk in S-RUTF was replaced with whey protein in A-RUTF.

The trial was limited by the large number of children who were lost to follow-up. Their outcomes were unknown; however, lost to follow-up or “default” was regarded as a negative outcome,
following other RUTF trials and international standards. Concerted efforts to seek malnourished children in Malawi who were lost to follow-up found that death or hospitalization occurred at about twice the rate as in those who reached a definitive outcome. There were no differences between the characteristics of children lost to follow-up compared with those who reached a definitive outcome in this study (data not shown).

We used ITT analyses, which are considered the strongest approach for randomized clinical trials to ensure unbiased comparisons among the treatment groups. If a per protocol analysis had been conducted on these SAM data from Ghana, recovery rates would have been 92% and 96% for A-RUTF and S-RUTF, respectively ($P<.05$). If we assume that half of the children lost to follow-up had a definitive negative outcome, then recovery rates for SAM would have been 79% and 87% for A-RUTF and S-RUTF, respectively ($P=.05$). This study is one of very few published clinical trials treating acute malnutrition in Ghana. Previously, a treatment trial of SAM in the Upper East region found a recovery rate of 71% (95% CI=68.0% to 76.0%) and default at 28% (95% CI=24.0% to 32.0%), results that are similar to our findings from Brong Ahafo. While we believe that our data support the conclusion that A-RUTF is inferior to S-RUTF, this conclusion is tempered by uncertainty from children defaulting.

### TABLE 4. Comparison of Outcomes Between Assigned Treatment Food for Ghanaian Children With Severe Acute Malnutrition and Moderate Acute Malnutrition

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Assigned A-RUTF</th>
<th>Assigned S-RUTF</th>
<th>$P$ Value$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All study participants</strong></td>
<td>n=642</td>
<td>n=628</td>
<td></td>
</tr>
<tr>
<td>Defaulted,b No. (%)</td>
<td>116 (18.1)</td>
<td>68 (10.8)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Died, No. (%)</td>
<td>1 (0.2)</td>
<td>4 (0.6)</td>
<td>.21</td>
</tr>
<tr>
<td>Recovered, No. (%)</td>
<td>516 (80.4)</td>
<td>554 (88.2)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Remained malnourished, No. (%)</td>
<td>9 (1.4)</td>
<td>2 (0.3)</td>
<td>.06</td>
</tr>
<tr>
<td>Rate of weight gain,$c$ g/kg/d, mean (SD)</td>
<td>1.88 (1.8)</td>
<td>2.04 (2.0)</td>
<td>.31</td>
</tr>
<tr>
<td>Rate of MUAC gain,$c$ mm/d, mean (SD)</td>
<td>0.16 (0.2)</td>
<td>0.18 (0.2)</td>
<td>.04</td>
</tr>
<tr>
<td><strong>Participants with SAM</strong></td>
<td>n=199</td>
<td>n=202</td>
<td></td>
</tr>
<tr>
<td>Defaulted,b No. (%)</td>
<td>60 (30.1)</td>
<td>41 (20.3)</td>
<td>.03</td>
</tr>
<tr>
<td>Died, No. (%)</td>
<td>1 (0.5)</td>
<td>3 (1.5)</td>
<td>.62</td>
</tr>
<tr>
<td>Recovered, No. (%)</td>
<td>130 (65.3)</td>
<td>156 (77.2)</td>
<td>.01</td>
</tr>
<tr>
<td>Remained malnourished, No. (%)</td>
<td>8 (4.0)</td>
<td>2 (1.0)</td>
<td>.06</td>
</tr>
<tr>
<td>Rate of weight gain,$c$ g/kg/d, mean (SD)</td>
<td>2.40 (2.4)</td>
<td>2.90 (2.6)</td>
<td>.04</td>
</tr>
<tr>
<td>Rate of MUAC gain,$c$ mm/d, mean (SD)</td>
<td>0.20 (0.2)</td>
<td>0.25 (0.2)</td>
<td>.047</td>
</tr>
<tr>
<td><strong>Participants with MAM</strong></td>
<td>n=443</td>
<td>n=426</td>
<td></td>
</tr>
<tr>
<td>Defaulted,b No. (%)</td>
<td>56 (12.6)</td>
<td>27 (6.3)</td>
<td>.002</td>
</tr>
<tr>
<td>Died, No. (%)</td>
<td>0 (0.0)</td>
<td>1 (0.2)</td>
<td>&gt;.99</td>
</tr>
<tr>
<td>Recovered, No. (%)</td>
<td>386 (87.1)</td>
<td>398 (93.4)</td>
<td>.003</td>
</tr>
<tr>
<td>Remained malnourished, No. (%)</td>
<td>1 (0.2)</td>
<td>0 (0)</td>
<td>&gt;.99</td>
</tr>
<tr>
<td>Rate of weight gain,$c$ g/kg/d, mean (SD)</td>
<td>1.66 (1.5)</td>
<td>1.61 (1.5)</td>
<td>.62</td>
</tr>
<tr>
<td>Rate of MUAC gain,$c$ mm/d, mean (SD)</td>
<td>0.13 (0.2)</td>
<td>0.14 (0.2)</td>
<td>.29</td>
</tr>
</tbody>
</table>

Abbreviations: A-RUTF, alternative ready-to-use therapeutic food; MAM, moderate acute malnutrition; MUAC, mid-upper arm circumference; SAM, severe acute malnutrition; S-RUTF, standard ready-to-use therapeutic food.

$^a$ Statistical comparisons made using Student’s t test for continuous parameters and Fisher’s exact test for categorical parameters.

$^b$ Defaulters were treated as unrecovered in the calculation of recovery rates.

$^c$ Calculated for the first 4 weeks of treatment.

Our data support that A-RUTF is inferior to S-RUTF, but the conclusion may be affected by defaulting.
Our findings showed that the children receiving A-RUTF enrolled on both SAM and MAM criteria were more likely to default, and we do not have the information to explain why that occurred. In response to a question asked of every caregiver on every return visit, only 3 caregivers overall remarked that their children did not like consuming the RUTF. Most defaulting occurred after 1 or 2 visits. Because of the randomized trial design, we conclude that defaulting is caused by A-RUTF, rather than coincidental circumstances.

Formal acceptability testing was conducted in a crossover design using the RUTFs daily for a week in MAM children. A-RUTF and S-RUTF showed similar amounts consumed, respectively (93% and 92%, \(P > .05\)) and similar liking scores attributed by the mother (3.5 and 4.1, \(P > .05\)), and there were no differences in adverse effects. It is possible children did not like the A-RUTF equivalently over a longer period of time, which led caregivers to not return; however, most defaulting was seen during the first few weeks of treatment. It is unlikely that the whey substituted for milk resulted in the inferior outcomes among the A-RUTF group because RUTF formulations have previously used whey with noninferior outcomes. The acceptability study results, timing of defaulting, and surveys of returning caregivers all indicate that organoleptic inferiority of A-RUTF did not result in the reduced recovery rate.

A study in Malawi found that a milk-free soya, maize, sorghum, amino acid-supplemented RUTF was not inferior to S-RUTF in treating children with SAM. This study was conducted in a more controlled environment with study participants returning daily to “daycare sites” for supervised feeding. A Zambian study used a similar sorghum RUTF without amino acids in a noninferiority trial and concluded that sorghum RUTF was inferior to S-RUTF in children. A possible explanation for...
The inferior outcomes among children receiving the dairy-free sorghum RUTF may have been the acknowledged inferior protein quality compared to S-RUTF. We found that a novel A-RUTF, which also included soy and sorghum, as well as having a similar protein quality as S-RUTF, caused less Ghanaian children to recover from SAM.

Under stressful physiological states, such as during rapid growth, nucleotides are required in the diet for optimal host response.28 There are limited data describing the nucleotide content in foods; purine tables are most frequently used to estimate nucleotide content. Grains, such as sorghum, have a low purine content,29,30 which suggests that A-RUTF had a lower nucleotide content than S-RUTF. However, we were not able to detect differences in nucleotide content between the foods in the untargeted metabolomics assays.

<table>
<thead>
<tr>
<th>Metabolite Class</th>
<th>Specific Metabolites Identified Only in Alternative RUTF</th>
<th>Pathobiological Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phosphatidylcholines</td>
<td>1,2-Dipalmitoleoyl-sn-glycero-3-phosphocholine</td>
<td>Major component of most biological membranes, found in soy foods</td>
</tr>
<tr>
<td></td>
<td>1-O-Hexadecenyl-2-deoxy-2-thio-5-acetyl-sn-glycero-3-phosphocholine</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-Stearyl-2-myristoyl-sn-glycero-3-phosphocholine</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-Palmitoyl-2-docosahexaenoyl-sn-glycero-3-phosphocholine</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PC(O:16:0/16:1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PC(18:1/20:2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PC(18:0/20:4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Palmitoyleicosapentaenoyl phosphatidylcholine</td>
<td></td>
</tr>
<tr>
<td>Phosphoethanolamines</td>
<td>2-Arachidonoyl-1-palmitoyl-sn-glycero-3-phosphoethanolamine</td>
<td>Ethanolamine derivative of phospholipids</td>
</tr>
<tr>
<td></td>
<td>2-Linoleoyl-1-palmitoyl-sn-glycero-3-phosphoethanolamine</td>
<td></td>
</tr>
<tr>
<td>Cholesterol</td>
<td>4-Cholestenone</td>
<td>Oxidized forms of cholesterol, the likely source in RUTF is dairy products</td>
</tr>
<tr>
<td></td>
<td>7-Oxocholesterol</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4-Beta-hydroxycholesterol 4-acetate</td>
<td></td>
</tr>
<tr>
<td>Ceramides</td>
<td>Ceramide (18:1/16:0)</td>
<td>A lipid component of cell membranes that enhances membrane rigidity and facilitates cell signaling through the membrane</td>
</tr>
<tr>
<td></td>
<td>N-Palmitoyl-o-sphingosine</td>
<td></td>
</tr>
<tr>
<td>Phytosterols</td>
<td>Cholestan-3-one</td>
<td>Plant-derived sterols typically found in soy products</td>
</tr>
<tr>
<td></td>
<td>Dihydroradaidzein</td>
<td></td>
</tr>
<tr>
<td>Isoflavones</td>
<td>Genistin</td>
<td>Isoflavonoid compounds almost entirely derived from legume species, interact with estrogen receptors</td>
</tr>
<tr>
<td></td>
<td>Glyctin</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6'-O-Acetylgenericin</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6'-O-Acetylglyctin</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Daidzin</td>
<td></td>
</tr>
<tr>
<td>Vitamins</td>
<td>Flavine mononucleotide</td>
<td>Form of riboflavin</td>
</tr>
<tr>
<td>Glucosyl glucose</td>
<td>(3beta,5xi,9xi,18xi,22beta)-22,25-Dihydroxystearic-12-en-3-yl 6-deoxy-alpha-L-mannopyranosyl-(1-&gt;2)-beta-o-glucopyranosiduronic acid</td>
<td>A small carbohydrate component of cellulose</td>
</tr>
<tr>
<td>Specific metabolites identified only in standard RUTF</td>
<td>1-Palmitoyl-2-stearoyl-sn-glycero-3-phosphocholine</td>
<td>Major component of most biological membranes</td>
</tr>
<tr>
<td></td>
<td>1-Docosahexaenoyl-2-stearoyl-sn-glycero-3-phosphocholine</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-Palmitoyl-2-oleoyl-sn-glycerol</td>
<td></td>
</tr>
<tr>
<td>Lipids</td>
<td>Erucic acid</td>
<td>Minor components of edible oils; fatty acid and a monoglyceride</td>
</tr>
<tr>
<td></td>
<td>Glycerol 1-myristate</td>
<td></td>
</tr>
<tr>
<td>Phenylpropanoids</td>
<td>14-(Methylpentadecanoylamino)-3-phenylpropanoic acid</td>
<td>Food additive made from cinnamic acid and a natural product in coffee and tea</td>
</tr>
<tr>
<td></td>
<td>3,5-Dimethoxy-4-hydroxyphenylacetic acid</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3-Hydroxy-4-methoxycinnamic acid</td>
<td></td>
</tr>
<tr>
<td>Phenylethylamide</td>
<td>Phenylethylamide 359</td>
<td>Flavoring agent, naturally occurs in peanut</td>
</tr>
</tbody>
</table>

**Table 5.** Untargeted Metabolomic Assessment of A-RUTF and S-RUTF

**Abbreviations:** A-RUTF, alternative ready-to-use therapeutic food; S-RUTF, standard ready-to-use therapeutic food.
Our cost data indicate for SAM that A-RUTF and S-RUTF are similar per child recovered, in spite of a 14% cost reduction per kilogram for the A-RUTF. No savings would be realized by using A-RUTF compared to S-RUTF in SAM.

While A-RUTF and S-RUTF met international specifications for nutrient content, the greater amounts of protein and fat in S-RUTF compared to A-RUTF led to greater rates of weight gain, but these seem unlikely to be important factors in increasing defaulting. RUTF specifications were determined on the basis of expert opinion, not clinical evidence; thus, protein and fat requirements may not be optimal. Some form of food intolerance may possibly have occurred with A-RUTF, resulting in greater default rates. This problem has been observed by the senior author in the past with RUTF made with chickpea in Africa.

With regard to the bioactive metabolites in A-RUTF compared to S-RUTF, the presence of isoflavonoids might have contributed to the poorer outcome. Isoflavonoids have metabolic effects to reduce lipogenesis, which is often thought to be an advantage for healthy consumers in the developed world; however, in this population of acutely malnourished children, this would not be the case. No xenobiotics or toxins were found in A-RUTF.

The sum of the evidence presented here indicates that A-RUTF is inferior to S-RUTF; it causes lower recovery in SAM and MAM, as well as lower rates of weight and MUAC gain in SAM. It is most important that RUTF facilitate recovery in SAM because SAM causes the most deaths. The certainty of this evidence is tempered by the observation that most failures in our trial were the result of defaulting, and the definitive outcomes in those cases are unknown. In conclusion, we cannot endorse this A-RUTF as noninferior to S-RUTF, and we recommend caution and further testing before any alternative RUTF is used in an operational setting. These data emphasize the utility of randomized trials to assess different RUTFs that meet international standards to determine equivalence.

### Funding:
Supported by the Children’s Investment Fund Foundation (CIFF) and the Hickey Family Foundation (HFF). The opinions expressed herein are those of the authors and do not necessarily reflect the views of CIFF or HFF.

### Competing Interests:
None declared.

### REFERENCES
17. FAO; IFAD; UNICEF; WFP; WHO. The State of Food Security and Nutrition in the World 2018: Building Climate Resilience for Food
Alternative RUTF Yields Lower SAM Recovery Than the Standard


© Kohlmann 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-19-00004

Peer Reviewed

Received: December 22, 2018; Accepted: April 2, 2019; First Published Online: June 12, 2019

Evaluating WHO-Recommended Interventions for Preterm Birth: A Mathematical Model of the Potential Reduction of Preterm Mortality in Sub-Saharan Africa

Jennifer B. Griffin,a Alan H. Jobe,b,c Doris Rouse,a Elizabeth M. McClure,a Robert L. Goldenberg,d Beena D. Kamath-Rayneb,c,e

Using the Maternal and Neonatal Directed Assessment of Technology (MANDATE) model, we estimate that WHO-recommended interventions could have saved nearly 300,000 lives in 2015. Combined interventions had the greatest impact. MANDATE can allow health officials to prioritize implementation strategies.

ABSTRACT

Background: Preterm birth, a leading cause of neonatal mortality, has the highest burden in low-income countries. In 2015, the World Health Organization (WHO) published recommendations for interventions to improve preterm outcomes. Our analysis uses the Maternal and Neonatal Directed Assessment of Technology (MANDATE) model to evaluate the potential effects that WHO-recommended interventions could have had on preterm mortality in sub-Saharan Africa in 2015.

Methods: We modeled preterm birth subconditions causing mortality (respiratory distress syndrome, intraventricular hemorrhage, necrotizing enterocolitis, sepsis, birth asphyxia, and low birth weight). For each subcondition, models were populated with estimates of WHO-recommended intervention prevalence, case fatality, coverage, and efficacy. Various scenarios modeled improved coverage of single and combined interventions compared with baseline.

Results: In 2015, approximately 500,000 neonatal deaths due to preterm birth occurred in sub-Saharan Africa. Single interventions with the greatest impact on preterm mortality included oxygen/continuous positive airway pressure (44,000 lives saved), cord care (38,500 lives saved), and breastfeeding (30,200 lives saved). Combined with improved diagnosis/transfer to a hospital, the impact of interventions showed greater reductions in mortality (oxygen/continuous positive airway pressure, 134,100 lives saved; antibiotics, 28,600 lives saved). Combined interventions had the greatest impact. Together, hospital delivery with comprehensive care for respiratory distress syndrome saved 190,600 lives, and comprehensive thermal care, breastfeeding, and prevention/treatment for sepsis saved 94,400 lives.

Conclusion: In 2015, WHO-recommended interventions could have saved the lives of nearly 300,000 infants born preterm in sub-Saharan Africa. Combined interventions are necessary to maximize impact. Mathematical models such as MANDATE can estimate effects on health outcomes to allow health officials to prioritize implementation strategies.

INTRODUCTION

Preterm birth is the largest cause of neonatal mortality worldwide, with best estimates of 15 million infants affected yearly. Rates appear to increase in countries as data systems improve.1–2 Complications of preterm birth result in significant risks for developmental disability in survivors and high costs for long-term complex health care needs.3 In 2012, the Born Too Soon report highlighted the problem by publishing country-specific rates of preterm birth and calling for implementation of simple interventions that decreased preterm birth complications in high-income countries prior to the influence of neonatal intensive care.3 A further complication is that the causes of preterm birth are multifactorial, and classification of a phenotype of preterm birth is imprecise because of heterogeneous clinical presentations and confounding factors such as maternal malnutrition and infections.3–6

With the understanding that innovative solutions are needed to decrease mortality from preterm birth, the World Health Organization (WHO) published recommendations in 2015 on interventions to improve
quality of care and outcomes surrounding preterm birth. The report detailed both maternal and neonatal interventions administered during pregnancy, labor, delivery, or the early neonatal period with the best available evidence for improving the incidence and adverse outcomes of preterm birth.

Given limited resources and the priorities of governmental agencies and national subgroups to implement the guidelines, our aim was to identify interventions with greater effects on improving mortality due to preterm birth. Such interventions would be the focus of initial efforts at implementation. We used the Maternal and Neonatal Directed Assessment of Technology (MANDATE) model to evaluate WHO-recommended interventions for preterm birth to determine which interventions and/or bundle of interventions had the most impact in terms of lives saved.

METHODS

We considered the WHO-recommended interventions that could be provided during pregnancy, labor, and the neonatal period for reducing neonatal mortality in preterm infants. Recognizing that this list is not exhaustive in terms of additional challenges faced by premature infants, such as asphyxia and sepsis, we sought other literature to determine the WHO-recommended interventions for these conditions as well. Because WHO did not focus on interventions associated with the prevention or reduction of preterm birth (e.g., pregestational agents), these were outside the scope of this analysis.

We used MANDATE, a nonstochastic, decision-tree model, to evaluate how WHO-recommended interventions would have influenced mortality in preterm neonates in sub-Saharan Africa in 2015. The methods used to develop the MANDATE model have been previously described. Briefly, we conducted a systematic review to populate variables regarding penetration, utilization, and efficacy of preventive, diagnostic, and treatment interventions specific to preterm birth in sub-Saharan Africa. Penetration, defined as the availability of an intervention, and utilization, defined as the appropriate use of an intervention, were considered in home, clinic, and hospital settings. Efficacy, defined as the ideal therapeutic effect of a given intervention, was treated as constant regardless of setting. MANDATE differentiates between the efficacy of diagnostics, which typically falls into 3 categories: (1) symptom recognition, made by a caregiver or unskilled care provider, frequently in a home setting; (2) clinical diagnostics, made by a skilled provider; and (3) technology-based diagnostics used to formally diagnose a condition.

We conducted the review, using PubMed, MEDLINE, the Cochrane Library, and WHO databases from 1980 to 2015 and the search terms “preterm,” “mortality OR death,” with “intervent* OR prevent* OR diagnost* OR treat*” and “developing countries OR low income countries OR sub-Saharan Africa.” For intervention efficacy parameters, we used a modified GRADE system to prioritize higher quality data. Demographic and Health Surveys, United Nations, and WHO data were used to populate key parameters regarding the number of births, prevalence of prematurity, and case fatality rate data. Parameters were reviewed by experts on preterm mortality in low-income countries and were incorporated into the model. Model building was an iterative process, with calibration against high-level WHO estimates. The version of MANDATE used for this analysis can be accessed at http://mnhtech.org.

Table 1 summarizes the interventions recommended by WHO related to preterm birth mortality, as well as MANDATE model assumptions regarding the penetration, utilization, and efficacy of these interventions in sub-Saharan Africa. WHO recommendations included quantification of the strength of recommendations (weak, strong, or conditional) by considering the quality of evidence (graded as very low, low, moderate, or high) and were up to date as of 2015, but they are expected to be updated as new data accrue. MANDATE assumes births occur across different settings, including the home (50%), clinics (35%), and hospitals (15%), based on the most recently available Demographic and Health Survey data for sub-Saharan African countries. Additionally, we assumed that chlorhexidine use would be appropriate for all home births across sub-Saharan Africa, based on the uncertainty bounds for UNICEF-estimated neonatal mortality rates for sub-Saharan Africa in 2015. These generalizing assumptions may not reflect regional, country, and local variation in birth rates and neonatal mortality rates. The online model allows MANDATE users to change these assumptions to reflect additional data that may be available for a specific region or country, or to reflect other differences (e.g., urban/rural).

For each subcondition affecting preterm mortality, the model was populated with the estimated prevalence, the case fatality rate, and WHO-recommended interventions to prevent
<table>
<thead>
<tr>
<th>Intervention</th>
<th>Recommendation Summary</th>
<th>WHO Strength of Recommendation for Implementation</th>
<th>Quality of Evidence</th>
<th>Baseline Penetration in MANDATE Home/Clinic/Hospital, %</th>
<th>Baseline Utilization in MANDATE Home/Clinic/Hospital, %</th>
<th>Efficacy in MANDATE Model, %</th>
<th>Key References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prenatal interventions for preterm</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antenatal corticosteroids</td>
<td>For women at risk of preterm birth (24–34 weeks gestation) under specific conditions</td>
<td>Strong</td>
<td>Moderate</td>
<td>0/10/50</td>
<td>0/5/25</td>
<td>RDS: 50</td>
<td>16–18, 28</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Antibiotics for pre-term labor</strong></td>
<td>For women with preterm prelabor rupture of membranes</td>
<td>Strong</td>
<td>Moderate</td>
<td>Not included in model</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Postnatal care</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cord care</td>
<td>Daily CHX application to the umbilicus for newborns born at home in settings with high neonatal mortality. Clean, dry cord care for newborns born in health facilities and at home in low neonatal mortality settings.</td>
<td>Strong</td>
<td>Moderate</td>
<td>0/0/0</td>
<td>0/0/0</td>
<td>55</td>
<td>29–32</td>
</tr>
<tr>
<td><strong>Care of the preterm/LBW neonate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thermal care for preterm newborns</td>
<td>KMC for the routine care of newborns weighing ≤2,000 g at birth, and should be initiated in health care facilities as soon as the newborns are clinically stable.</td>
<td>Strong</td>
<td>Moderate</td>
<td>95/95/95</td>
<td>0/0/2</td>
<td>51</td>
<td>15, 33</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unstable newborns weighing ≤2,000 g or stable newborns weighing ≤2,000 g who cannot be given KMC should be cared for in a thermoneutral environment either under radiant warmers or in incubators.</td>
<td>Strong</td>
<td>Very low</td>
<td>0/0/50</td>
<td>0/0/30</td>
<td>60</td>
<td>34–36</td>
</tr>
<tr>
<td>Feeding</td>
<td>LBW infants, including those with very low birth weight, should be fed mother’s own milk.</td>
<td>Strong</td>
<td>Moderate</td>
<td>99/99/99</td>
<td>20/40/55</td>
<td>Sepsis: 55</td>
<td>37–39</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LBW: 18</td>
<td></td>
</tr>
<tr>
<td>Management: newborn resuscitation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate drying and additional stimulation</td>
<td>Newly born babies who do not breathe spontaneously after thorough drying should be stimulated by rubbing the back 2–3 times before cord clamping and PPV initiation.</td>
<td>Weak</td>
<td>Not graded</td>
<td>50/85/90</td>
<td>50/70/85</td>
<td>15</td>
<td>40–42</td>
</tr>
<tr>
<td>PPV</td>
<td>In newly born term or preterm (&gt;32 weeks of gestation) babies requiring PPV, ventilation should be initiated with air.</td>
<td>Strong</td>
<td>Moderate</td>
<td>5/50/95</td>
<td>20/40/60</td>
<td>40</td>
<td>42–45</td>
</tr>
<tr>
<td>Oxygen therapy for preterm newborns</td>
<td>Ventilation of preterm babies born at or before 32 weeks of gestation with oxygen therapy with 30% oxygen or air (if blended oxygen is not available).</td>
<td>Strong</td>
<td>Very low</td>
<td>0/15/60</td>
<td>0/50/75</td>
<td>RDS: 25</td>
<td>46, 47</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Asphyxia: 25</td>
<td></td>
</tr>
<tr>
<td>Management: RDS</td>
<td>Continuous positive airway pressure for newborns with RDS</td>
<td>Strong</td>
<td>Low</td>
<td>0/2/20</td>
<td>0/50/70</td>
<td>RDS: 50</td>
<td>46, 47</td>
</tr>
<tr>
<td></td>
<td>Continuous positive airway pressure therapy is recommended for the treatment of preterm newborns with RDS.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Asphyxia: 50</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 1. WHO Interventions and Recommendations to Improve Preterm Birth Mortality, With MANDATE Model Assumptions of Intervention Penetration, Utilization, and Efficacy in Sub-Saharan Africa, 2015**
preterm mortality (Table 2). Subconditions included direct causes of preterm mortality, including respiratory distress syndrome (RDS), intraventricular hemorrhage (IVH), and necrotizing enterocolitis (NEC), sepsis, birth asphyxia, and low birth weight. Each scenario models 2 levels of improvement: (1) an incremental care model, in which penetration and utilization are increased by 20% from current care estimates, with maximum penetration and utilization set at 98%; and (2) a universal coverage model, in which penetration and utilization are set to 98% for each intervention.

Baseline MANDATE estimates of the number of neonatal preterm deaths associated with subconditions impacting preterm mortality in sub-Saharan Africa in 2015 can be seen in the Figure. The methods describing the calculation of baseline mortality estimates have been previously published. In short, we began with sub-Saharan African pregnancies in 2015; calibrated the model using historic rates for the various conditions (Table 2); estimated the impact of interventions using penetration, utilization, and efficacy (Table 1); and applied untreated case fatality rates (Table 2) in order to estimate baseline mortality. Direct complications of preterm birth contributing to preterm mortality, including RDS, IVH, and NEC, were estimated to cause 303,400 deaths, consistent with other reports of preterm deaths due to direct complications in sub-Saharan Africa. Other subconditions associated with preterm mortality included sepsis and birth asphyxia. Finally, prematurity alone posed an increased risk of death among preterm neonates. These preterm babies at risk of mortality were captured in the model subcondition “low

### TABLE 1. Continued

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Recommendation Summary</th>
<th>WHO Strength of Recommendation for Implementation</th>
<th>Quality of Evidence</th>
<th>Baseline Penetration in MANDATE Home/Clinic/Hospital, %</th>
<th>Baseline Utilization in MANDATE Home/Clinic/Hospital, %</th>
<th>Efficacy in MANDATE Model, %</th>
<th>Key References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surfactant administration for newborns with RDS</td>
<td>Surfactant replacement therapy is recommended for intubated and ventilated newborns with RDS.</td>
<td>Conditional (health care facilities only with intubation, ventilator care, blood gas analysis, newborn nursing care and monitoring)</td>
<td>Moderate</td>
<td>0/1/5</td>
<td>0/50/75</td>
<td>35</td>
<td>46,48</td>
</tr>
</tbody>
</table>

**Management: neonatal sepsis**

| Prophylactic antibiotics for prevention of sepsis | A neonate with risk factors for infection (i.e., membranes ruptured > 18 hours before delivery, maternal fever > 38°C before delivery or during labor, or foul-smelling or purulent amniotic fluid) should be treated with the prophylactic antibiotics ampicillin and gentamicin for at least 2 days and reassessed if signs of sepsis or positive blood culture. | Weak | Very low | Not modeled |

| Empirical antibiotics for suspected neonatal sepsis | Neonates with signs of sepsis should be treated with antibiotic treatment for at least 10 days. | Strong | Low | 10/85/95 | 20/65/75 | 72 | 49,50 |

**Management: NEC**

| Antibiotics for treatment of NEC | Young neonates with suspected NEC should be treated with intravenous or intramuscular ampicillin (or penicillin) and gentamicin as first-line antibiotic treatment for 10 days. | Strong | Low | Not modeled |

Abbreviations: CHX, chlorhexidine; KMC, kangaroo mother care; IVH, intraventricular hemorrhage; LBW, low birth weight; MANDATE, Maternal and Neonatal Directed Assessment of Technology; NEC, necrotizing enterocolitis; PPV, positive pressure ventilation; RDS, respiratory distress syndrome; WHO, World Health Organization.
MANDATE estimated approximately 500,000 total preterm deaths in 2015 associated with direct and indirect conditions that contribute to preterm mortality in sub-Saharan Africa. We modeled the impact of WHO-recommended interventions for each subcondition potentially causing preterm death and summarized scenarios for each subcondition and associated interventions impacting preterm mortality. All scenarios are compared with the baseline, current-care scenario. In most cases, results are rounded to the nearest 100 preterm deaths.

RESULTS

In Table 3, we report the number of deaths from preterm-associated mortality from RDS, IVH, and NEC in sub-Saharan Africa in 2015. With interventions at current levels of use, approximately 300,000 preterm deaths can be attributed to RDS, IVH, and NEC (scenario 1). In the first set of WHO-recommended interventions, even in the universal coverage models, only small to moderate impacts on preterm mortality were present, with 300 lives saved with improved surfactant use in hospitals (scenario 2), 5,000 lives saved in the antenatal corticosteroid model (scenario 3), and 42,300 lives saved in the improved oxygen/continuous positive airway pressure (CPAP) model (scenario 4). Thus, there was a decrease in mortality of approximately 14%. The second set of scenarios evaluated improved diagnosis with and without transfer. Improved diagnosis of preterm labor alone had a smaller impact (scenario 5).
compared with diagnosis with transfer to a higher-level facility (scenario 6), with 2,100 lives saved compared with baseline in the incremental change model and 16,300 lives saved in the universal coverage model. Combining improved diagnostics and transfer to a higher-level care with single interventions (scenarios 7 and 8) demonstrated the synergistic effects of improving diagnostics, transfers, and treatments. For example, improved diagnosis of RDS paired with transfer and improved CPAP prevented 16,000 preterm deaths in the incremental change model and 127,300 deaths in the universal coverage model, a 42% reduction. In the final set of scenarios, we show that incremental and near universal improvements in diagnosis and transfer with WHO-packaged interventions would have the greatest impact on preterm mortality. For example, improved preterm labor diagnosis, transfer, antenatal corticosteroids, surfactant, and oxygen/CPAP jointly prevented approximately 112,000 preterm deaths (scenario 10) and improved diagnosis of respiratory distress, transfer, surfactants in hospitals, and oxygen/CPAP have prevented 155,700 preterm deaths in the universal coverage model, a reduction of nearly half (scenario 11). In the model with the greatest impact, scenario 12, all preterm deliveries were assumed to occur in hospital settings, with improved antenatal corticosteroid use and diagnosis and treatment of respiratory distress, including surfactants and oxygen/CPAP, thereby preventing the deaths of 190,600 preterm infants in this universal coverage model, a mortality reduction of nearly two-thirds.

In Table 4, we report the number of deaths from preterm-associated mortality from sepsis, birth asphyxia, and low birth weight in sub-Saharan Africa in 2015. We estimated preterm mortality associated with sepsis, birth asphyxia, and low birth weight with current levels of care, including low to moderate coverage of positive pressure ventilation (PPV), oxygen, cord care, breastfeeding, and antibiotics, to be approximately 198,000 deaths (scenario 1). In the first set of improved WHO single intervention scenarios, we found small to moderate impacts on preterm mortality, with near-universal oxygen/CPAP for birth asphyxia saving 1,700 lives (scenario 2); PPV for birth asphyxia saving 4,200 lives (scenario 3); drying and stimulation of newborns saving 3,000 lives (scenario 4); thermal care, including kangaroo mother care (KMC), saving 9,100 lives (scenario 5); antibiotics saving 18,200 lives (scenario 6); breastfeeding, reducing mortality from
both sepsis and LBW, saving 30,200 lives (scenario 7); and chlorhexidine in home settings and dry cord care in clinical settings saving 38,500 lives, a reduction of nearly 20% (scenario 8). The second set of scenarios shows that improved diagnosis and transfer on preterm mortality from sepsis, birth asphyxia, and low birth weight had relatively small impacts on preterm mortality, with a range from diagnosis of birth asphyxia saving 1,900 lives (scenario 9) to sepsis diagnosis with transfer to clinical settings saving 14,300 lives (scenario 13). The third set of scenarios demonstrates the increased impact of improved diagnosis and transfer with WHO-recommended single interventions, with diagnosis and transfer for birth asphyxia with oxygen support saving 6,800 lives (scenario 13);
<table>
<thead>
<tr>
<th>Scenario No.</th>
<th>Scenario</th>
<th>Incremental Change Model</th>
<th>Universal Coverage Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Preterm Deaths, No. d</td>
<td>Preterm Deaths Prevented Compared With Current Level of Care, No. (%)</td>
</tr>
<tr>
<td>1</td>
<td>Current levels of prevention, diagnosis, and treatment</td>
<td>198,400</td>
<td>N/A</td>
</tr>
<tr>
<td>2</td>
<td>Improved WHO single interventions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Oxygen/CPAP for birth asphyxia in clinics and hospitals</td>
<td>198,000</td>
<td>400 (0.2)</td>
</tr>
<tr>
<td>3</td>
<td>PPV for birth asphyxia in all settings</td>
<td>197,200</td>
<td>1,200 (0.6)</td>
</tr>
<tr>
<td>4</td>
<td>Drying and stimulation for birth asphyxia in all settings</td>
<td>196,486</td>
<td>1,900 (1.0)</td>
</tr>
<tr>
<td>5</td>
<td>Thermal care for LBW, including KMC in all settings and warmers in hospital settings</td>
<td>196,000</td>
<td>2,500 (1.3)</td>
</tr>
<tr>
<td>6</td>
<td>Antibiotics for suspected neonatal sepsis in all settings</td>
<td>192,100</td>
<td>6,300 (3.2)</td>
</tr>
<tr>
<td>7</td>
<td>Breastfeeding for sepsis and LBW in all settings</td>
<td>189,300</td>
<td>9,100 (4.6)</td>
</tr>
<tr>
<td>8</td>
<td>Chlorhexidine for sepsis in home settings and dry cord care in clinical settings</td>
<td>190,800</td>
<td>7,600 (3.8)</td>
</tr>
<tr>
<td>9</td>
<td>Improved diagnosis and transfer with current care</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Diagnosis of birth asphyxia and need for postresuscitation care, with current levels of care</td>
<td>197,200</td>
<td>1,300 (0.7)</td>
</tr>
<tr>
<td>10</td>
<td>Diagnosis of birth asphyxia and need for postresuscitation care, with improved transfer to hospitals, with current levels of care</td>
<td>197,000</td>
<td>1,400 (0.7)</td>
</tr>
<tr>
<td>11</td>
<td>Diagnosis of sepsis, with current levels of care</td>
<td>194,700</td>
<td>3,700 (1.9)</td>
</tr>
<tr>
<td>12</td>
<td>Diagnosis of sepsis and transfer to hospitals, with current levels of care</td>
<td>187,400</td>
<td>11,000 (5.5)</td>
</tr>
<tr>
<td>13</td>
<td>Improved diagnosis and transfer with WHO single treatment interventions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Diagnosis of birth asphyxia and need for postresuscitation care, transfer, and oxygen/CPAP</td>
<td>196,300</td>
<td>2,100 (1.1)</td>
</tr>
<tr>
<td>14</td>
<td>Diagnosis of birth asphyxia and need for postresuscitation care, transfer, and positive pressure ventilation</td>
<td>195,500</td>
<td>2,900 (1.5)</td>
</tr>
<tr>
<td>15</td>
<td>Diagnosis of sepsis, transfer, and antibiotics for suspected neonatal sepsis</td>
<td>180,800</td>
<td>17,600 (8.9)</td>
</tr>
<tr>
<td>16</td>
<td>Improved diagnosis and transfer with WHO-packaged interventions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Drying and stimulation, diagnosis of birth asphyxia and need for postresuscitation care, transfer to hospitals, and treatment, including PPV and oxygen/CPAP</td>
<td>188,057</td>
<td>10,400 (5.2)</td>
</tr>
<tr>
<td>17</td>
<td>Cord care and breastfeeding, diagnosis of sepsis, transfer, and antibiotics for suspected neonatal sepsis</td>
<td>169,200</td>
<td>29,200 (14.7)</td>
</tr>
<tr>
<td>18</td>
<td>Packaged interventions 16 and 17, with increased thermal care and breastfeeding for LBW</td>
<td>159,300</td>
<td>39,100 (19.7)</td>
</tr>
</tbody>
</table>

Abbreviations: CPAP, continuous positive airway pressure; KMC, kangaroo mother care; LBW, low birth weight; PPV, positive pressure ventilation; WHO, World Health Organization.

a Assumptions regarding baseline penetration and utilization of interventions including ANCS, surfactant, and CPAP as shown in Table 1. Assumptions regarding diagnostics and transfers found in Table 2.
b The incremental change model assumes 20% increase from baseline penetration and utilization.
c The universal coverage model assumes 98% penetration and utilization of interventions.
d All estimates rounded to nearest 100.
DIAGNOSIS OF BIRTH ASPHYXIA WITH TRANSFER AND PPV SAVING 8,600 LIVES (SCENARIO 14); AND IMPROVED SEPSIS DIAGNOSIS AND TRANSFER WITH ANTIBIOTICS SAVING 28,600 LIVES (SCENARIO 15). IN THE LAST SET OF SCENARIOS, COMBINED INTERVENTIONS HAD THE GREATEST IMPACT ON PRETERM MORTALITY. IMPACTS RANGED FROM 26,200 LIVES SAVED WITH COMPREHENSIVE CARE FOR BIRTH ASPHYXIA (SCENARIO 16); TO 59,100 LIVES SAVED WITH COMPREHENSIVE PREVENTION AND TREATMENT OF SEPSIS; TO 94,400 LIVES SAVED WITH THERMAL CARE AND BREASTFEEDING (SCENARIO 17), IN ADDITION TO COMPREHENSIVE CARE FOR BIRTH ASPHYXIA AND SEPSIS (SCENARIO 18).

DISCUSSION

TO IMPROVE NEONATAL MORTALITY WORLDWIDE, THE BURDEN OF PRETERM-RELATED DEATHS MUST BE LESSENED. THE BORN TOO SOON REPORT SETS A TARGET OF 50% REDUCTION IN PRETERM DEATHS IN COUNTRIES WITH A NEONATAL MORTALITY RATE ABOVE 5 PER 1,000 LIVE BIRTHS BY 2025. TOWARD THIS END, CLOSING THE GAP ON THE HIGHER INCIDENCE MORTALITY FROM PRETERM BIRTH AND ITS SUBSEQUENT COMPLICATIONS, SPECIFICALLY IN LOW- AND MIDDLE-INCOME COUNTRIES, IS ONE OF THE TOP PRIORITIES. TO HELP POLICY MAKERS AND FRONTLINE HEALTH PROVIDERS, THIS STUDY SHOWCASES THE SINGLE INTERVENTIONS OR BUNDLES OF INTERVENTIONS RECOMMENDED BY WHO THAT COULD POTENTIALLY HAVE THE GREATEST EFFECT ON REDUCING PRETERM BIRTH MORTALITY.

THE SINGLE INTERVENTIONS WITH THE GREATEST IMPACT ON PRETERM MORTALITY ARE OXYGEN/CPAP, CORD CARE, BREASTFEEDING, AND ANTIBIOTICS. INTERESTINGLY, IN A LIVES SAVED TOOL (LiST) ANALYSIS OF PRETERM BIRTH INTERVENTIONS, INCLUDING FAMILY PLANNING, ANTENATAL CORTICOSTEROIDS, ANTIBIOTICS FOR PROLONGED PREMATURE RUPTURE OF MEMBRANES, IMMEDIATE ASSESSMENT AND SIMPLE CARE OF ALL BABIES, NEONATAL RESUSCITATION, THERMAL CARE, AND KMC, 84% OF PREMATURE BABIES COULD BE SAVED IF THESE INTERVENTIONS WERE MADE UNIVERSALLY AVAILABLE (95%).

In that study, the 2 single interventions that had the greatest impact on preterm birth mortality were antenatal corticosteroids and KMC. The terms of our analysis were different. We modeled lives saved for Sub-Saharan Africa for a single year (2015), while the LiST analysis considered lives saved for 2 separate periods of time (2010–2015) and then through 2025. Furthermore, the LiST analysis did not include the single intervention that had the greatest effect in our analysis: oxygen/CPAP for RDS. Our model assumptions were also different; MANDATE assumes that treatments are relevant only to preterm neonates with/without particular subconditions. For example, we assumed that KMC will only be efficacious in stable preterm neonates without critical subconditions needing treatment, such as RDS or sepsis, while LiST assumes that KMC is relevant to the entire population of preterm neonates. We also assumed that receipt of treatment interventions is dependent on a previous diagnosis, while LiST modeled the combined effect of multiple interventions. Finally, the MANDATE model provides the opportunity to evaluate the location where the intervention is implemented and account for transfer to higher levels of care; for example, it is unlikely that oxygen could be implemented in a home setting, but is likely to be used in a hospital setting.

The benefits of antenatal corticosteroids in low-resource settings are unclear. A recent WHO multicountry survey on maternal and newborn health indicated that current national coverage estimates varied, between 16% and 91%, with a median of 54%. A multicountry cluster-randomized trial done by the National Institute of Child Health and Human Development (NICHD) Global Network, the ACT trial, demonstrated increased 28-day neonatal mortality possibly explained by maternal infection in the corticosteroid-exposed group. Furthermore, in the ACT trial, only 16% of women who were given corticosteroids gave birth to an infant below the 5th percentile for weight, indicating unnecessary overexposure to the treatment.

Secondary analysis of the data from the Guatemala site suggested that the combination of improved quality of obstetric and neonatal care in facilities associated with antenatal corticosteroid treatment may have reduced neonatal mortality. A randomized controlled trial of late preterm infants (34 weeks to 36 weeks, 5 days) by the NICHD Maternal Fetal Medicine Unit Network demonstrated modest improvement in neonatal respiratory morbidity in older preterm infants with antenatal corticosteroid exposure, but an increased risk of hypoglycemia. Given concerns about safety and efficacy of antenatal corticosteroids in low-resource settings, WHO has strict criteria for their use, including accurate assessment of gestational age, imminent preterm birth, no clinical evidence of maternal infection, adequate childbirth care (recognition and management of preterm labor and birth), and adequate care for preterm newborn (including resuscitation, thermal care, feeding support, infection treatment, and safe oxygen use). An international research collaboration called WHO ACTION (Antenatal Corticosteroids for Improving Outcomes in Preterm Newborns) is conducting 2 concurrent placebo-controlled efficacy trials of antenatal...
Combined interventions together with transfer to the hospital had the greatest impact on lives saved.

corticosteroids (dexamethasone) that will eventually enroll over 28,000 women. For our analysis, we have assumed benefit only for antenatal corticosteroids based on the magnitude of effect in the multiple randomized controlled trials, primarily performed in high-income countries.

Combined interventions together with transfer to the hospital had the greatest impact on lives saved. This outcome is likely because hospitals have greater availability of some of the recommended interventions and are more likely to use an intervention, if available. However, the model assumes that interventions are being utilized at a gold standard level, which frequently will not be accurate. A study in India determined that while rates of institutional deliveries in South and Central India increased, and perinatal and stillbirth mortality decreased, neonatal mortality did not change. A 10-country analysis of skilled birth attendance demonstrated that less than 10% of mothers who saw a skilled birth attendant once during pregnancy received a set of 8 key interventions, and quality of care did not increase as number of antenatal visits increased. Furthermore, it is important to recognize that even when interventions are utilized, they may not change outcomes. For example, a recent cluster randomized study evaluated the use of the WHO Safe Childbirth Checklist accompanied by 8 months of coaching in facilities across India. While there was increased adherence to the 18 essential birth practices in the intervention group, maternal and perinatal mortality did not differ between groups. The study exemplifies the difficulty in achieving the high reliability of gold standard performance, even for the most evidence-based practices; for example, birth attendants performed hand hygiene in 35% cases in intervention groups. Clearly, complex relationships exist between quality of care and outcomes that are beyond the scope of what the MANDATE model, or in fact any model, can predict.

Our study has several other limitations. The assumptions in the model are based upon our best efforts to find primary sources that document the effects of interventions in low- and middle-income countries; therefore, the list of interventions included in our study is not exhaustive. While MANDATE has the ability to look with more granularity at some of the subconditions and the locations where interventions may occur, the lack of primary source data to populate our assumptions is an identified gap, which underscores the need for greater documentation and research in these areas. Due to varied quality of primary sources from a range of countries across sub-Saharan Africa, MANDATE makes the simplifying assumption that baseline condition incidence and intervention penetration and utilization are the same across the sub-Saharan Africa continent. MANDATE is a nonstochastic model decision tree model that does not model uncertainty. Furthermore, we could not distinguish between the effects of separate interventions for different degrees of prematurity (e.g., extremely preterm versus late preterm infants) or the interaction of multiple pathologies that may lead to a poor outcome. In some locales with extremely limited resources, we were also unable to determine if there was a threshold of prematurity below which interventions would be counterproductive because the chance of intact survival is so low. Finally, our study was not able to evaluate the possible adverse consequence of broader coverage of some interventions, for example, unnecessary treatment with antenatal corticosteroids or inappropriate overuse of oxygen in preterm infants, which could result in excess morbidity or mortality.

Moving forward, reducing preterm birth mortality requires improving coverage of evidence-based interventions that are known to reduce preterm birth-associated mortality, but then carefully quantifying the effects. The WHO guideline also described steps toward successful dissemination and implementation of the recommended interventions. However, the introduction of evidence-based policies to improve preterm birth outcomes depends on well-planned and participatory consensus-driven processes of adaptation and implementation. Evidence-based approaches to facilitate this knowledge and transfer exchange include collective impact collaboratives that bring together multiple sectors to achieve policy change through a common agenda, shared measurement systems, mutually reinforcing activities, continuous communication, and the presence of a backbone organization, as well as learning collaboratives that bring policy makers together in an ongoing way to share knowledge about a specific health outcome. To address issues of implementation and dissemination, further national and subnational groups will need support to adapt and implement the WHO-recommended interventions and change the beliefs and behaviors of local health care providers. The published recommendations specifically document anticipated barriers to implementation and possible steps to mitigate these challenges. Some of the anticipated barriers included nonavailability or an irregular supply of essential medicines,
lack of human resources with expertise and skill to implement recommended practices and monitor response, low certainty of gestational age estimation in low-resource environments, and lack of effective referral mechanisms and care that ensure management of women with preterm labor and preterm infants occurs within a continuum of care.

Finally, while the WHO-recommended guidelines represent the best available evidence-based interventions to decrease preterm birth mortality, further research is urgently needed for preterm birth prevention. A recent study indicated that while the strongest individual risk factor of preterm birth is previous preterm birth and preeclampsia, more than 65% of the total aggregated risk of preterm birth lacked a plausible biologic explanation. In addition, 63% of the differences in prematurity rates between countries could not be explained with known factors. New efforts to better classify the characteristics of the preterm birth syndrome, its clinical phenotypes, and core outcomes for evaluation of interventions will aid in focusing and accelerating research on this complicated topic.

Funding: This research was funded by a grant from the Bill & Melinda Gates Foundation to RTI International. The Gates Foundation did not play a role in the collection, analysis, or presentation of these data.

Competing Interests: None declared.

REFERENCES


47. PATH. Intrapartum-related events rapid landscape analysis. 2012. CrossRef. Medline


Introduction of Community-Based Provision of Subcutaneous Depot Medroxyprogesterone Acetate (DMPA-SC) in Benin: Programmatic Results

Tishina Okegbe,a Jean Affo,b Florence Djihoun,b Alexis Zannou,b Odilon Hounyo,b Gaston Ahounou,c Karamatou Adegnika Bangbola,c Nancy HARRISd

Lay community health workers and facility-based health care providers in Benin were trained to administer DMPA-SC safely and effectively in 10 health zones. Community-based DMPA-SC was popular, particularly among new users of contraception, and could help the country achieve its family planning goals.

Résumé en français à la fin de l’article.

ABSTRACT

The Republic of Benin faces high maternal, newborn, and child mortality; low modern contraceptive use; and a critical shortage of health workers. In 2013, the Government of Benin made 3 reproductive health commitments to improve national health indicators, including expanding provision of family planning services at the community level through task sharing. Since 2016, the Advancing Partners & Communities (APC) project has been helping the Benin Ministry of Health (MOH) provide subcutaneous depot medroxyprogesterone acetate (DMPA-SC; brand name Sayana Press) through facility-based health care providers and community health workers known as relais communautaires (RCs). DMPA-SC is an easy-to-administer, discreet injectable contraceptive that provides 3 months of protection from pregnancy. Beginning in May 2017, the government introduced DMPA-SC through a phased approach in 10 health zones, which encompassed 149 health centers and 614 villages. Between June 2017 and June 2018, the MOH and APC trained 278 facility-based providers and 917 RCs to provide DMPA-SC, and nearly 11,000 doses were subsequently administered to 7,997 women at facilities and in communities. This article presents findings from an assessment of community-level and health facility service data collected during the first 13 months of DMPA-SC introduction in Benin. Because of this intervention, nearly 35,000 women received family planning counseling and 7,997 women chose DMPA-SC. At the community level, 3,111 DMPA-SC users were first-time users of modern contraception. The initial success of the DMPA-SC rollout in Benin shows promise for helping the country meet its reproductive health commitments.

INTRODUCTION

Historically, the Republic of Benin has had low family planning use. Desire for large families and myths and misconceptions about family planning contribute to a low modern contraceptive prevalence rate of 12% and a high total fertility rate of 5.7.1 Yet while conservative social norms still hinder family planning use, the modern contraceptive prevalence rate has doubled since 2006, when it was 6.1%. There has been no corresponding or significant decrease in the total fertility rate, which stood at 6.0 in 1996.2 The percentage of adolescents under 19 years of age who have begun bearing children has declined only slightly since 1996, from 26% to 20% in 2018.1,2 In 2018, nearly 1 in 3 women (32%) wanted to limit or space her next birth but was not using a family planning method.1 The current modern family planning method mix is dominated by implants (45.0%), followed by injectables (18.3%).1

Expanding access to contraception—especially in communities that have an unmet need and demand for it—is essential to reach the country’s family planning goals. At the 2013 International Conference on Family Planning in Addis Ababa, the Government of Benin (GOB) made 3 commitments to improve reproductive health indicators.3 These commitments, which acknowledged the need to improve geographic access to family planning services...
planning services for people who did not have it, were to increase the national modern contraceptive prevalence rate from 7.9% to 20% by 2018; offer free family planning services and commodities to adolescents by 2019; and introduce family planning at all levels of the health system, including injectable contraception at the community level. As part of its commitment to the Ouagadougou Partnership, the GOB pledged to enable 2.2 million additional women to use modern family planning by 2020.

With an expanding population of 11 million, nearly half of whom reside in rural areas with limited access to health care services, Benin continues to face high maternal and child mortality. This problem is compounded by the country’s critical shortage of health workers. Other countries with similar shortages have introduced task-sharing approaches, as recommended by the World Health Organization (WHO), in which highly skilled health workers in facilities delegate less complex tasks to health workers who have fewer skills, thus better distributing the workload.

In 2014, the Benin Ministry of Health’s (MOH’s) Maternal and Child Health Directorate approved a pilot project to allow aides-soignantes to administer the 2-month intramuscular injectable Noristerat in Adja-Ouére commune in southeastern Benin. Aides-soignantes are a cadre of paraprofessional nurses’ aids who are stationed at health centers and provide immunization and assist in health promotion campaigns in communities. Previously, contraceptives were available exclusively at health facilities, which limited rural women’s access to family planning. During the 9-month pilot, aides-soignantes counseled 1,809 women on a range of voluntary family planning methods, and 449 women adopted Noristerat. The aides-soignantes referred 249 women to health centers for other methods including pills, intrauterine devices, and implants. Most clients had a positive response to the community-based delivery of family planning services. An assessment of the pilot concluded that community health workers (CHWs) can administer injectable contraceptives safely and effectively in Benin, and this evidence was the foundation for scaling up the approach.

**PROJECT DESCRIPTION**

Since 2012, the Advancing Partners & Communities (APC) project, funded by the United States Agency for International Development (USAID), has worked in more than 40 countries to advance and support programs that improve the overall health of communities and other health-related indicators, especially related to family planning. In Benin, between 2015 and 2018, the APC project built the technical and organizational capacity of 3 local NGOs—Dedras, Bupdos, and Sia N’son—to train relais communautaires (RCs), the lowest-level CHW cadre, to provide high-quality community-level primary health services including family planning in 10 USAID priority health zones. In 2016, the MOH authorized the introduction of DMPA-SC in health facilities and via RCs. Midwives administered DMPA-SC at health facilities, adding to their routine family planning service provision. RCs selected for training were linked to NGOs that received APC technical assistance.

DMPA-SC has been described as a “game-changing” injectable contraceptive because it is discreet and requires minimal training to administer. It is packaged as a prefilled, 3-month, all-in-one device and has the potential to greatly expand contraceptive access to women in need, particularly those in underserved areas. Recent studies have shown that CHWs can safely and acceptably administer DMPA-SC in communities. Studies have also shown that women can be trained to self-inject,

**BOX 1. Relais Communautaires in Benin**

Relais communautaires (RCs) constitute a key component of the national community health strategy. They typically receive supervision from staff at public health facilities and training and small monthly stipends from local NGOs. RCs are volunteers chosen by residents of their villages to provide a package of high-impact health services (paquet d’interventions à haut impact) related to water, sanitation, and hygiene; maternal and child health; nutrition; malaria; and family planning (condoms and resupply of oral contraceptive pills). RCs are required to be proficient in French. An estimated 15,000 are deployed nationally and 2,080 work in the United States Agency for International Development’s 10 priority health zones. The Ministry of Health offers RC refresher trainings every 2–3 years.

Benin has 2 types of RCs: those who provide curative services and those who provide preventive services only. Preventive RCs conduct health education sessions in their communities, while curative RCs also provide basic health care interventions and refer community members who need complex care to a local health facility. According to the national community health strategy, 1 curative RC should serve between 30 and 50 households that are 5 kilometers or more from the nearest public health facility. In practice, however, there are not enough trained RCs to fulfill this ideal.
where the practice is permitted and the product is registered as a self-injectable, which allows greater contraceptive control over their fertility.\textsuperscript{11,12}

After authorization for DMPA-SC introduction was granted, Beninese stakeholders, including the MOH and implementing partner staff, visited DMPA-SC pilot sites in Burkina Faso and Uganda to learn from their introduction processes (Box 2). These visits illustrated the benefits of task sharing for family planning and laid the groundwork for further investments in community-based family planning in Benin. Based on what the Beninese stakeholders learned, it was decided that DMPA-SC introduction would start in the 6 health zones in which intramuscular depot medroxyprogesterone acetate (DMPA-IM) was available (i.e., where clients and health care providers were already familiar with DMPA). Further, targeted communication would be developed in those areas; specific monitoring and evaluation tools for DMPA-SC would be developed; and a pharmacovigilance strategy would be established. Finally, before the pilot started, national and local political and administrative authorities would be engaged.

A national technical steering committee provided overall guidance and technical leadership for the introduction and implementation of DMPA-SC. The committee comprised technical and financial partners including APC, the United Nations Population Fund, USAID, the Benin Association for Social Marketing and Communication for Health (ABMS), the Beninese Association for Family Planning, and University Research Company’s Advancing Newborn, Child, and Reproductive Health project. The committee adapted the training curriculum, job aids, and monitoring and evaluation tools, all in French, that PATH developed for DMPA-SC introduction in Senegal.

The 3-day RC training curriculum included a review of all available family planning methods (types, uses, side effects, and eligibility for and advantages of use); how to conduct family planning counseling and referrals; and how to complete data collection tools. The process of adapting and formatting the tools for Benin took 9 months and several committee meetings to solicit input and ensure that content adhered to national health policies, strategies, and norms. For example, the committee decided that DMPA-SC would be injected only in the back of a client’s upper arm, rather than in the thigh or abdomen because of concerns for client privacy with male providers. (Pfizer’s Sayana Press product is labeled for administration in either the thigh or abdomen; however, the product has been shown to be equally effective when injected in the back of the upper arm.\textsuperscript{13} The WHO endorses injecting DMPA-SC in the upper arm and acknowledges that this injection site may be more comfortable for some women.) Once adapted, a field test with selected RCs and health care providers in Savalou-Banté health zone was conducted to determine the usability of the materials. Participant feedback was incorporated into the final versions.

DMPA-SC was brought into Benin under a waiver granted by the National Drug Authority in early 2017 and was registered in July 2017. The GOB, guided by the committee, introduced DMPA-SC through public-sector health facilities, and RCs linked to those facilities used the existing infrastructure instead of an NGO-led parallel supply chain. The process required integrating DMPA-SC into national health system tools, reporting forms, and the District Health Information System 2 (DHIS 2). Beginning in May 2017, introduction proceeded in a phased rollout in 10 USAID priority health zones where RCs had been trained on family planning methods and counseling through the MOH’s \textit{paquet d’interventions à haut impact} (PIHI) curriculum (Figure 1). DMPA-SC introduction began in Abomey-Calavi health zone, followed by

**BOX 2. Guiding Lessons From PATH’s 4-Country DMPA-SC Introduction**

1. Introduce DMPA-SC in health zones where DMPA-IM is already being offered—because providers and clients are already familiar with DMPA—including eligibility criteria, reinjection timeline, and side effect profiles.
2. Develop targeted, regional communication materials in accordance with local requirements and manufacturer guidelines.
3. Develop specific monitoring and evaluation tools for DMPA-SC.
4. Establish an appropriate pharmacovigilance strategy.
5. Engage national and local political and administrative authorities early on to garner buy-in.

Abbreviations: DMPA-IM, intramuscular depot medroxyprogesterone acetate; DMPA-SC, subcutaneous depot medroxyprogesterone acetate.

Health worker training was cascaded (Figure 2) to foster local ownership; increase sustainability, efficiency, and rapport between facility- and community-level health workers; and decrease costs. Technical steering committee members conducted trainings for MOH staff, who served as national master trainers. Master trainers then trained health zone staff (health center heads and midwives), who then trained RCs. Each training class of 20 RCs had 2 days of theory and a 1-day practicum at a local health center. RCs learned about all available family planning methods, including DMPA-SC, and how to refer clients who had side effects or wanted longer-acting methods. Participants practiced counseling sessions in French and the local language and gave injections to salt-filled condoms. As part of certification requirements, within 4 weeks of the training RCs

Abbreviation: DMPA-SC, subcutaneous depot medroxyprogesterone acetate.
were required to administer 5 supervised injections to clients who chose DMPA-SC.

After certification to provide community-based DMPA-SC, RCs received several forms of support. PIHI NGOs and RC supervisors at local health facilities held monthly group supervision sessions at which RCs discussed challenges, successes, and lessons with their peers and were reminded to follow up with clients for reinjection after 3 months. Health center in-charges also conducted quarterly on-site supervision visits with RCs. Lastly, supervision teams of staff from APC, the MOH, ABMS, and the district health zone and health center in-charges visited RCs on site 4 to 6 weeks post training to provide support and resolve challenges.

DMPA-SC is delivered to health centers through the national supply chain, which operates a “pull” system, whereby each level orders commodities; products flow from the central medical stores to district warehouses to health centers. Upon certification, health centers give RCs 5 free doses of DMPA-SC. RCs obtain product resupply at health centers at the nationally set cost of 200 CFA (~US$0.35) per dose and resell it to clients for the same price. RCs are provided with safety boxes to dispose used DMPA-SC devices, which are discarded per medical waste disposal guidelines at the health center during routine monthly group supervision sessions.

ABMS and APC led community sensitization and national advocacy sessions to raise awareness of family planning, specifically DMPA-SC. Sensitization sessions in each introduction health zone included local elected officials and traditional and religious leaders, and featured short films on the demographic dividend, maternal mortality, and community-based injectables programming, followed by discussions. These workshops prepared traditional and political leaders to support RC DMPA-SC administration. A total of 536 local officials and 578 traditional and religious leaders participated in family planning sensitization and advocacy sessions. At the national level, ABMS and APC led briefings to inform parliamentarians and journalists of the benefits of family planning and the GOB’s commitment to make DMPA-SC widely available through RCs.

The APC project assessed community-level and health facility service data collected by RCs and midwives between June 2017 and June 2018, including number of DMPA-SC doses administered, DMPA-SC uptake, and number of women receiving family planning counseling. Results and lessons from the first 13 months of DMPA-SC introduction are described below.

**RESULTS**

DMPA-SC was introduced in 10 health zones in Benin through a phased approach that allowed
course corrections during rollout to 149 health centers and 614 villages. The 278 trainers (health center heads and midwives) and 917 RCs who were trained administered nearly 11,000 doses of DMPA-SC over the 13-month period. RCs administered 1,309 doses of DMPA-SC to women as reinjections, and facility staff administered 1,403 doses, for a total of 2,712 reinjections. Few clients reported side effects.

Between June 2017 and June 2018, 9,296 women at the health facility level and 24,947 women at the community level received family planning counseling. A total of 7,997 women—including those who were adopting a modern family planning method for the first time and those switching from other family planning methods—chose DMPA-SC. DMPA-SC uptake at the community level climbed steadily during the first year of introduction and peaked in June 2018. At the facility level, DMPA-SC uptake increased from September 2017 onward, peaking at nearly 600 new users in June 2018.

At the community level, 3,111 women were first-time users of modern contraception, and 769 women switched to DMPA-SC from another contraceptive method (Figure 3). Further, Figure 3 shows that of the 5 family planning methods available at the community level through RCs (DMPA-SC, pills, CycleBeads, condoms, and spermicide), DMPA-SC was the most popular method for current users of family planning to switch to, representing nearly 54% of all switchers. It was not possible to disaggregate how many women adopting DMPA-SC at the facility level were new users of family planning versus switchers. Figure 4 shows that for much of the year, overall DMPA-SC uptake at the community level was comparable with facility-level uptake.

During the same time frame, Noristerat uptake at the health facility level was high, with 13,698 women adopting the method, followed by DMPA-SC (4,117 women) and DMPA-IM (972 women) (Figure 5). Noristerat is the most popular injectable contraceptive choice in Benin, so it is unsurprising that there were fewer users of DMPA-IM and DMPA-SC during this period. However, the goal of DMPA-SC introduction is not to supplant current options; it is to expand client access and method choice.

Although stock-outs were not measured systematically, anecdotal reports suggest they occurred at both the facility and community levels during the first months of DMPA-SC introduction. In facilities, stock-outs likely occurred because staff had to adapt to their new and additional responsibility for reordering DMPA-SC from district warehouses. In communities, either RCs did not realize that they should seek DMPA-SC resupply from the health center or when they did seek it, health centers were out of the product. It is important to note that RCs must contend with frequent stock-outs of products related to other public

Abbreviation: DMPA-SC, subcutaneous depot medroxyprogesterone acetate.
sector health interventions because Benin still struggles to maintain its national health commodity supply chain.

**DISCUSSION**

DMPA-SC was introduced through the public sector at health centers and in communities to expand family planning access to and method choice for women. RCs administered 49% of the total DMPA-SC doses to women who chose the method. Furthermore, 80% of DMPA-SC doses at the community level were administered to first-time family planning users, suggesting that community-based distribution of DMPA-SC by RCs is an effective way to increase use of family planning. Although there were challenges throughout the process, the phased introduction allowed course corrections along the way, providing useful lessons for future DMPA-SC programming.

Community-based distribution of injectables via CHWs is a promising way to expand access to family planning services and commodities, particularly for women who live in remote areas. Our results are in line with findings from DMPA-SC introduction in Burkina Faso, Niger, Senegal, and Uganda, which demonstrated that community-based distribution of DMPA-SC by RCs is an effective way to increase use of family planning. Although there were challenges throughout the process, the phased introduction allowed course corrections along the way, providing useful lessons for future DMPA-SC programming.
based distribution of DMPA-SC has the potential to reach new acceptors of family planning as well as youth and adolescents.8–10

The introduction of DMPA-SC corresponded with an increase in the number of women receiving family planning counseling at both the community and facility levels. As this experience shows, adding a contraceptive option to the method mix often prompts programs to provide refresher trainings for providers to strengthen skills, thus enhancing the overall quality of the family planning program. Many RCs reported being happy with their new DMPA-SC-related responsibilities because they were providing a useful service to community members and were eager to continue.

Lessons Learned
The lessons learned during DMPA-SC expansion in Benin were used to improve subsequent trainings and strengthen implementation and can provide guidance to countries introducing DMPA-SC (Table).

<table>
<thead>
<tr>
<th>What Was the Challenge?</th>
<th>Why Was This a Challenge?</th>
<th>How Was the Challenge Addressed?</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC certification process took longer than anticipated</td>
<td>The process may take longer than expected for several reasons, including low client uptake for the new method, which prevents RCs from demonstrating competence. In addition, RCs may need additional support after initial training to manage their new duties (counseling and filling out data collection tools).</td>
<td>Build in “extra” time for RC certification because several factors can affect the process. Invite health zone coordinating doctors to attend supportive supervision visits to oversee health center staff who supervise RCs. (Note that current global guidance states that in countries where CHWs already administer injectables, certification can be granted with 3 injections)</td>
</tr>
<tr>
<td>Community awareness about DMPA-SC was low in regions where introduction occurred</td>
<td>Communication campaign was not launched simultaneously with the introduction of DMPA-SC.</td>
<td>Introduce a targeted communication campaign in regions where DMPA-SC is introduced to garner interest and create awareness about the method.</td>
</tr>
<tr>
<td>Stockouts of DMPA-SC at the facility and community level were frequent at the beginning of introduction</td>
<td>Health facility staff incorrectly assumed that the APC project was responsible for reordering DMPA-SC stock.</td>
<td>Clearly outline health facility manager responsibilities during training and encourage staff to take ownership for ordering commodities. At the national level, it is critical to ensure that the method appears on the national resupply order form as soon as the new method becomes available.</td>
</tr>
<tr>
<td>RC family planning counseling skills were low</td>
<td>RCs previously had been trained on family planning through MOH PIHI curriculum and initially APC only provided a short counseling skills session during the DMPA-SC training.</td>
<td>Include a half-day refresher training session on counseling skills so that all RCs have an opportunity to practice and enhance their counseling skills. A counseling job aid was also added to guide RCs through the counseling process and ensure informed choice and voluntarism.</td>
</tr>
<tr>
<td>Data collection was not standardized across RCs or health facilities</td>
<td>During the first few months of DMPA-SC introduction, RCs did not receive family planning registers (which were to be provided by their linked NGO) and therefore did not collect data in a systematic way.</td>
<td>During subsequent RC trainings, the importance of using the appropriate forms to collect client data was emphasized. Additionally, time was allotted for RCs to practice filling in the client family planning registers. At the health facility level, technical assistance is provided to providers during routine supervision visits to ensure that they are correctly collecting and reporting family planning data.</td>
</tr>
</tbody>
</table>

Abbreviations: APC, Advancing Partners & Communities; CHW, community health worker; DMPA-SC, subcutaneous depot medroxyprogesterone acetate; MOH, Ministry of Health; PIHI, paquet d’interventions à haut impact (package of high-impact health services); RC, relais communautaires (lay community health workers).
Build Facilitation Skills in CHW Trainers

Master trainers realized that many RC trainers had little experience facilitating sessions, which resulted in a low-quality training experience for RCs. The MOH and APC added a module on how to conduct effective trainings to the RC trainer curriculum. This half-day training increased RC trainer competency and led to higher-quality RC training sessions.

Pay Attention to Training Needs and Be Able to Provide Additional Support

During the post training supervision phase, many RCs displayed weak family planning counseling skills. To remedy this problem, the APC project developed a job aid that gave concise guidance on how to counsel on all family planning methods and ensure informed choice and voluntarism. This job aid was given to trained RCs and incorporated into the training curriculum. A half-day counseling refresher session was added to subsequent trainings to supplement the family planning information that RCs received under the government’s PIHI curriculum. The refresher training included time to review the job aid and practice its contents through role play. Lastly, a day of on-site counseling capacity building for RCs who continued to exhibit poor counseling skills post training and certification was added to MOH, APC, and ABMS routine monitoring and supervision visit schedule.

Build in Ample Time to Certify Community Health Workers

RC certification took longer than expected, sometimes exceeding the 4-week validation period initially established. Of the 917 RCs trained, approximately half were certified to administer DMPA-SC by June 2018. Those who did not become certified within 2 months of training are preventive RCs, conducting family planning and DMPA-SC awareness-raising activities in their communities. One cause of delayed certification was supervisors’ lack of adequate follow-up with RCs. In response, APC invited health zone coordinating doctors, who oversee health center staff, to attend the supportive supervision visits so they could note challenges, develop monitoring and action plans, and encourage supervisors to be available to RCs.

Recognize That Appropriate and Simultaneous Education and Communication Initiatives Are Critical to Family Planning Uptake

Another cause of delayed RC certification was low client uptake of DMPA-SC, which prevented RCs from administering the 5 voluntary client injections to demonstrate competence. Low client uptake of an unfamiliar (i.e., new) family planning option is common. To rectify this, RCs strengthened their mobilization efforts to challenge social norms that hinder family planning uptake. Client demand was also low, likely due to the delayed launch of a mass media DMPA-SC educational campaign. It is critical to launch corresponding communication campaigns (in accordance with local requirements and manufacturer guidelines) and product introduction simultaneously to create momentum and attract clients. The campaign should include information on the benefits and risks of family planning in general and the new method in particular, as well as where to access it.

National health worker strikes, which took place during RC training, also delayed RC certification. Facility-level health care providers serve as RC supervisors, so when they were unavailable during the strike, RCs were unable to complete the requisite number of client injections for certification. Many trained RCs were unable to provide community-based services during this period, either because they were not yet properly certified to do so, were not being directly supervised, or were not able to resupply stock.

Provide Sufficient Training to Enable Facility-Based Providers to Order Commodities Early and Ensure That MOH Includes New Commodities on Resupply Forms at the Onset of Product Introduction

Product stock-outs were initially problematic in the first health zone because health center staff were unaware of their responsibility for reordering commodities. The reason for this is likely twofold: (1) health center staff assumed that because APC was leading the DMPA-SC trainings, the project was also responsible for product resupply; and (2) the national commodities resupply form was not updated to include DMPA-SC until June 2017. Stock-outs at the facility level trickled down and affected community-level provision because RCs sought DMPA-SC resupply at health centers.

A continuing challenge is the collection of high-quality data. RCs record community-level data in family planning registers, which are

Collection of high-quality data is a continuing challenge, and more data on side effects experienced by clients are needed.
compiled at the facility level. During the early phases of implementation, some RCs did not have the registers because of delayed printing by the affiliated NGOs and therefore did not collect or record data in a standardized manner. Once those RCs received their registers, APC and MOH staff provided technical support during monitoring visits to ensure data were collected properly. Subsequent trainings emphasized the importance of monitoring, evaluation, and data collection and allowed time for RCs to practice completing the data registers. Data quality also remains a problem at the health facility level because reporting is primarily paper-based, which often results in poor record-keeping. APC and MOH staff provide technical assistance to facility-level health care providers during routine supervision visits to ensure that family planning data are captured and reported accurately. High-quality data are critical for measuring the impact of DMPA-SC introduction in the health zones.

Another ongoing challenge is limited data on client side effects. Women rarely reported side effects to RCs or facility-level health care providers, and although it is possible that this reflects a low rate of experienced side effects, it is also possible that clients fail to report side effects that they consider “minor” or health workers are not collecting or reporting this information accurately. We suggest establishing a more robust system to monitor and capture potential minor and major side effects.

Our final lesson is the importance of involving communities and their gatekeepers. We found that inviting local stakeholders—including religious and traditional leaders, local and national politicians, and communities at large—to orientation sessions during the DMPA-SC introduction process was critical to gaining their support, without which the program would not have succeeded.

Limitations
A limitation of this analysis is the lack of unique client identifiers within the data. Consequently, it is impossible to track individual women over time to determine if they elected to be reinjected with DMPA-SC and if so how many times; therefore, calculating continuation and discontinuation rates is also impossible. The indicators simply captured the total number of reinjections in communities and at health facilities during the specified period. Moreover, this type of routine service statistics data does not measure the number of contraceptive discontinuers, only users, so it is difficult to assess if and how much the program increased contraceptive use and contributed to national reproductive health goals. Further, because this analysis is restricted to service statistics data, client motivations for using family planning for the first time or choosing to switch to DMPA-SC from another method cannot be determined.

Additionally, given the MOH’s current facility-level indicators, it is impossible to determine if women adopting DMPA-SC were new users to family planning or had switched from another method. Having this disaggregated information in the future will help determine user profiles for women who choose DMPA-SC. The data show that in health centers where DMPA-SC and DMPA-IM were both offered, DMPA-SC was the more popular method, surpassing the number of DMPA-IM users during each of the 13 assessed months. It would be helpful to monitor from which methods users were switching to determine if DMPA-SC introduction decreases use of other methods, including DMPA-IM and Noristerat. Further, age-disaggregated data are not collected at either the facility or community level because the indicator has yet to be added to national data collection forms. This information is critical for improving youth-focused programming.

CONCLUSIONS
Results from the first 13 months of DMPA-SC introduction in Benin demonstrate the demand for community-based family planning provision, as uptake of the injectable contraceptive increased steadily at the community level. Between June 2017 and June 2018, 7,997 women chose to adopt DMPA-SC as their preferred method of family planning, from either a community- or facility-based health worker. Results suggest that service provision through RCs is feasible and acceptable among clients, as nearly 4,000 women—nearly 80% of whom were first-time users of modern contraception—chose to receive this method at the community level.

As the popularity of DMPA-SC grows in Benin, service provision through additional channels, including self-injection, should be considered to make the method available to all who would choose it. Recent findings from a DMPA-SC self-injection trial in Malawi show a user discontinuation rate of 56% when CHWs administer injections. However, when women self-inject, the discontinuation rate drops by almost half, to 30%. Although DMPA-SC self-injection is not
Introduction de l’offre communautaire du Sous-cutané Depot Medroxyprogesterone Acetate (DMPA-SC) au Bénin : résultats programmatiques

Les agents de santé communautaires et les prestataires de soins de santé des formations sanitaires publiques au Bénin ont été formés pour adminiser le DMPA-SC de manière sûre et efficace dans 10 zones sanitaires. Le DMPA-SC à base communautaire est populaire, particulièrement Alison Davis et Fortuné Challa. Special acknowledgement would also like to thank USAID/Benin for their ongoing support, Pius Gounadon and partner NGOs Dedras, Sia N’son, and Bupolas. We would also like to thank USAID/Benin for their ongoing support, particularly Alison Davis and Fortuné Challa. Special acknowledgement goes to Kristen Devlin for her careful manuscript review and helpful feedback.

Acknowledgments: We would like to thank the Benin Ministry of Health for their collaboration in implementing this project, particularly Drs. Didier Agossadou, Franck Robert Zammou, Aïhanasse Houmanankan and Pius Gounadon and partner NGOs Dedras, Sia N’son, and Bupolas. We would also like to thank USAID/Benin for their ongoing support, particularly Alison Davis and Fortuné Challa. Special acknowledgement goes to Kristen Devlin for her careful manuscript review and helpful feedback.

Competing Interests: None declared.

REFERENCES
en particulier parmi les nouveaux utilisateurs de contraception, et peut aider le pays à atteindre ses objectifs en matière de planification familiale.

RÉSUMÉ

La République du Bénin est confrontée à une forte mortalité maternelle, néonatale et infantile; faible utilisation de contraceptifs modernes; et une pénurie critique d’agents de santé. En 2013, le gouvernement du Bénin a pris 3 engagements en matière de santé reproductive afin d’améliorer les indicateurs de santé nationaux, notamment dans l’expansion de la fourniture de services de planification familiale au niveau communautaire par le biais d’une délégation des tâches. Depuis 2016, le projet Advancing Partners & Communities (APC) aide le Ministère béninois de la Santé (MS) à fournir de l’acétate de médroxyprogesterone-dépôt sous-cutané (DMPA-SC; nom de marque Sayana Press) par l’intermédiaire des prestataires de soins de santé des formations sanitaires et d’agents de santé communautaires connus sous le nom de relais communautaires (RC). Le DMPA-SC est un contraceptif injectable discret, facile à administrer, qui offre une protection contre la grossesse pendant 3 mois. À partir de mai 2017, le gouvernement a introduit le DMPA-SC selon une approche progressive dans 10 zones sanitaires comprenant 149 centres de santé et 614 villages. Entre juin 2017 et juin 2018, le MS et APC ont formé 278 prestataires des formations sanitaires et 917 RC à fournir du DMPA-SC, et près de 11 000 doses ont ensuite été administrées à 7 997 femmes dans les formations sanitaires et dans les communautés. Cet article présente les résultats d’une évaluation des données sur les services au niveau communautaire et des formations sanitaires collectées au cours des 13 premiers mois de l’introduction du DMPA-SC au Bénin. Grâce à cette intervention, près de 35 000 femmes ont bénéficié de conseils en matière de planification familiale et 7 997 d’entre elles ont choisi d’utiliser DMPA-SC. Au niveau communautaire, 3 111 utilisatrices du DMPA-SC utilisaient pour la première fois la contraception moderne. Le succès initial du déploiement du DMPA-SC au Bénin montre qu’il est prometteur d’aider le pays à respecter ses engagements en matière de santé de la reproduction.

Peer Reviewed

Received: December 20, 2018; Accepted: April 9, 2019; First Published Online: June 6, 2019


© Okegbe 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-19-00002
Are Procured Quantities of Implants Adequate and Appropriate? Modeling Procurement, Inventory, and Consumption of Contraceptive Implants During Rapid Uptake

Laila Akhlaghi,a Alexis Heaton,a Yasmin Chandani

Recent rapid increases in implant procurement have not resulted in system overstocks to date. We found no standard factor for relating inventory quantities to consumption rates. Rather, that relationship requires specific understanding of the country supply chain, inventory control parameters, and current and future demand.

ABSTRACT
Donors and others are concerned that implants procured under the Family Planning 2020 Initiative exceed the number sought by clients, resulting in accumulating stocks. To explore this issue, we examined 3 questions across 9 countries: (1) How accurate were procurement quantities given requirements for filling supply chains for the rapidly growing implant programs? (2) Is there a standard factor that can be applied to consumption data to predict procurement volumes required? (3) How accurately do demographic estimates mirror dispensed-to-client data? We created a model incorporating public-sector supply chain system parameters to calculate system "imputed" inventory and the system "filled-to-max" inventory. Comparing results determined the adequacy of the procurement quantities. The proportion of consumption that the filled-to-max inventory represented through time suggests whether a standard factor can be applied to consumption to predict necessary procurement volumes. We compared demographic estimates to consumption data to determine the usability of the former in predicting demand. According to model results, 3 of the 9 countries came close to procuring accurate quantities over the study period between 2010 and 2017, 4 had procurement volumes lower than what was required to fill the supply chain to maximum inventory requirement levels, and 2 had volumes that exceeded the need. We found no standard factor for relating inventory quantities to consumption rates across countries, given that inventory needs can vary based on system design parameters and the rates of growth or decline in consumption. Finally, we observed that our demographic estimates were on average lower than the dispensed-to-client data in the 6 countries for which these data were available. Study results show that the significant investments in procurement quantities for the rapidly growing implant programs were justified based on consumption and system design. This research should assure observers that rapid increases in implant procurement quantities (where data are available) have generally not resulted in overstocks of the system to date. It suggests that the relationship between procurement quantities and consumption levels cannot be accurately assessed without understanding the country supply chain, inventory control parameters, and current and future demand.

BACKGROUND
The Family Planning 2020 (FP2020) initiative was launched in 2012 in recognition of high unmet need for contraception, especially in low- and middle-income countries. FP2020 set a target of reaching an additional 120 million women with contraception by the year 2020.1 An important part of the strategy to achieve the goal of increased modern contraceptive prevalence rate (mCPR) has been increasing access to long-acting reversible contraceptives, specifically hormonal contraceptive implants.

Despite their high level of effectiveness, initial uptake of contraceptive implants was low, due to the high upfront purchase price of products and barriers associated with introduction.2,3 Recognizing these limitations, a consortium of public and private organizations—the Bill & Melinda Gates Foundation; the Clinton Health Access Initiative; the governments of Norway, Sweden, the United Kingdom, and the United States; and the Children’s Investment Fund Foundation, with support from the United Nations Population Fund-collaborated to establish the Implants Access Program. Through this program, Bayer Pharma’s Jadelle and Merck/MSD’s Implanon and Implanon NXT were made available to women in the world’s poorest countries, as recognized by FP2020, through price reductions of approximately 50%4 from 2013 through 2018, backed by volume guarantees ensuring that minimum quantities would be purchased each year.
As a result of the price reductions and efforts to increase accessibility via more trained service providers, procurements of contraceptive implants in the FP2020 markets and their consumption have risen substantially since 2013. Implants as a proportion of total mCPR continue to increase in many countries, as measured by surveys such as Demographic and Health Surveys (DHS) and Performance Monitoring and Accountability 2020 (PMA2020).7,8

The rapid uptake of implants combined with limited data visibility into consumption in many countries generated questions from donors, procurers, and other partners about the relationship between product volumes procured and product volumes consumed—specifically, they wanted to know whether increases in procurement quantities were resulting in increases in implant use. Initial review of procurement quantities and demographic surveys, especially in a number of the high-volume procuring countries, led to concerns that procurement quantities far exceeded the number of women seeking this method and were resulting in stocks accumulating in countries,9 and the objectives of the Implants Access Program were consequently not being met.

While it was hypothesized that some of the excess quantities were procured for inventory, the lack of access at the global level to country supply chain data (logistics management information systems) meant that this hypothesis could not be studied effectively across countries using readily available survey statistics. To further explore the relationship between procurement volumes and consumption, we reviewed logistics and demographic data from 9 countries to understand 3 questions:

1. How accurate were procurement quantities given requirements for filling supply chains for the rapidly growing implant programs?
2. Is there a standard factor that can be applied to consumption data to predict procurement volumes required?
3. How accurately do demographic estimates mirror dispensed-to-client data?

**METHODS**

In order to determine whether procurement quantities were accurate for the inventory needs of countries, we created an Excel-based model whereby we could input supply chain system parameters as defined by the country’s supply chain design, historic procurements (quantities entering the supply chain), and estimated consumption and/or implants dispensed to clients (quantities leaving the supply chain via patient use). The model used these data to calculate the imputed inventory for each month of the review period. Figure 1 depicts the imputed inventory for the country example of Ethiopia, and the details on the calculation steps are provided in Supplement 1.

Using the same model, we built a second scenario using the same estimates of consumption, but instead determined the system filled-to-max inventory required to fill the system to maximum stock level for a given country’s supply chain design. Figure 2 depicts the filled-to-max inventory for Ethiopia as an example.

A comparison of the 2 results (imputed and filled-to-max inventory) shows either a surplus or deficit for each month of the model as shown in Figure 3.

Using the following equation, we could determine the percentage difference for each month (formula below) and answer our first question about accuracy of procurement quantities.

\[
\text{Percentage difference} = \frac{\text{imputed inventory} - \text{filled-to-max inventory}}{\text{filled-to-max inventory}}
\]

The result gives us an estimate of procurement quantity accuracy by month. If there was no difference between the imputed and the filled-to-max inventory quantities (i.e., a percentage difference of 0%), procurement quantities for the system were deemed to be accurate and appropriate for that month. If the percentage difference was above 0%, it would indicate overprocurement, and if the percentage difference was below 0%, the indication would be underprocurement.

With the results of the filled-to-max inventory, we could then calculate the proportion of the consumption that the filled-to-max inventory represented through time. The resulting ratio answered our second question about a standard factor. Calculation steps for answering the second research question are included in Supplement 2.

Because we collected 2 different inputs for demand or consumption, the first being dispensed-to-client data reported by information systems in country (where available) and the second estimated using demographic data, the 2 were compared to determine their relationship and answer our third question about the accuracy of demographic estimates.
As part of the model design, we chose to input the data and capture outputs by month through the review period rather than use annual aggregates and averages. Most of our inputs were already available as monthly figures such as those for dispensed-to-client data and procurements. In addition, multiple PMA2020 reports may exist per given year. We also assumed that monthly results...
would better align with actual movements in the supply chains because reporting and distribution in many settings is done on a monthly, bimonthly, quarterly, or other periodic basis depending on the level of the supply chain. The only input data not available on a monthly schedule were survey results, so best efforts were made to use midpoint months when analyzing demographic surveys instead of publication dates. Linear interpolation was used to determine monthly estimates between points.

Finally, we evaluated the model outputs of imputed inventory by comparing the model results with actual collected and reported system inventory from 1 country, as discussed in more detail in the Model Evaluation section.

Defining Procurement, Inventory, and Supply Chain Parameters

For most FP2020 countries, the national supply chain for imported contraceptives, including implants, typically begins at the port of entry. Once products clear customs, they are typically received at the central warehouse, distributed to regional and/or district distribution points (and potentially other sublevels), and used to resupply health facilities, commonly known as service delivery points (SDPs). In this article, we refer to this flow as the in-country supply chain. Contraceptives need to flow through each level of this supply chain to be available as inventory at the SDP as a choice for a woman seeking contraception on any given day. The more levels in an in-country supply chain, the longer the time needed for products to reach the SDP and the further in advance they have to be procured.

Countries determine procurement quantities based on supply (or procurement) plans using forecasted demand. Forecasts should ideally be based on historic demand and growth patterns, with demographic parameters being a less ideal foundation or an alternative method used for comparison. Adjustments to forecasts can be made based on market intelligence of trends and/or other factors assumed to affect future demand (such as budgets, goals, trainings, and/or policies). Supply plans determine quantities procured by taking into account this forecasted demand, available inventory, inventory control policies at each level of the supply chain (such as maximum and minimum stocking parameters), storage capacity, and potential loss/expiry.

We evaluated the model accuracy by comparing its imputed results with collected and reported system inventory for 1 country.
Holding inventory at each level of the supply chain is critical to buffer against uncertainty of demand (demand variability), as forecasts are often inaccurate. Additional inventory (or safety stock) is particularly important for a product with increasing consumption, such as contraceptive implants, for which it is difficult to anticipate rapid changes in demand. Additional inventory (shelf-life allowing) is also especially critical in the public sector supply chains for many of the countries in the FP2020 market where procurements may happen only once or twice a year and the ability to respond quickly to sudden changes in demand or resources is limited.

To ensure sufficient quantities of products when and where they are needed, the entire supply chain should be filled with inventory in advance of demand and refilled routinely. As products are given to clients or distributed from one point to another, inventory held at higher levels will flow down to the next level (under most demand variabilities) to make sure that product is available to meet future client demand at SDPs. Ensuring that sufficient inventory is available at all levels of the supply chain requires accurate forecasts, routine and accurate data on stock from all levels of the system, and timely procurements. As part of supply chain design, supply chain systems establish inventory control parameters to determine how much of a product needs to be held at each level based on frequency and lead time for resupply. Typically, the inventory control parameters are reflected in a relative measure of months of stock and use current and projected consumption levels per month.

For example, a system that requires 18 months of stock to fill the supply chain to maximum system levels (e.g., 5 months at central level + 4 months at regional level + 4 months at district level + 4 months at SDP + 1 month for the community health worker) and using 15,000 units a month would require about 270,000 implants to meet client demand. For a similar system that sees demand increase about 75% from month 1 to month 18 (15,000 units in month 1 to more than 26,000 units in month 18) would require 370,000 units (100,000 more units for the full period of months 1–18). A waterfall graph (Figure 4) depicts the quantity needed to be procured (to enter the system) and flow through the various distribution points to reach the point of consumption.

**FIGURE 4.** Pictorial Depiction of Maximum Inventory Needs (System and Level) and Flow Through In-Country Supply Chain With Growing Demand in a Sample Country

Ensuring that sufficient inventory is available requires accurate forecasts, routine and accurate data on stock, and timely procurements.
Data
To use the model developed to determine if procurement quantities are adequate and appropriate to meet the demand of women choosing implants, we reviewed data available from 9 countries (Burkina Faso, Ethiopia, Ghana, Kenya, Pakistan, Tanzania, Uganda, Country A, and Country B). Country selections were based on availability of (1) procurement data (described further below); (2) demographic data (at least 2 time points), listed in the Appendix, used to estimate users of implants; (3) a reliable source of dispensed-to-client consumption (Table 1), a second alternative figure for estimating users of implants; and (4) information on the in-country supply chain design (Table 2). Two of the 9 countries have been anonymized (Country A and Country B) to enable inclusion of their data in this article.

Procurement and Beginning Balance
Through the Implants Access Program, the research team collected monthly procurement data for all FP2020 countries for levonorgestrel 2-rod, 5-year implant and etonogestrel 1-rod, 3-year implant. Although Sino-implant was procured during this time, available data (from the Reproductive Health Interchange) indicate that procurement volumes were either through social marketing organizations dispensing to markets we did not account for (Ethiopia [DKT] and Pakistan [Marie Stopes International]) or were at percentages below 1% of total procurement (Burkina Faso, Ghana, Kenya, and Uganda).

The procurement data provided by the manufacturers capture the dates of shipments from the manufacturing point but do not include the date of arrival or clearance in the destination country. To obtain this additional information, we reviewed country reports (e.g., procurement plans captured in databases using PipeLine, a software commonly used for supply planning and monitoring) to identify receipt dates for shipments into countries. If these data were not available in the country supply plan, as is frequently the case for NGO and social marketing organization orders, an estimate was made based on the available country-specific estimated lead times for shipping, evidenced by the above country reports.

Baseline inventory holdings were estimated at the start, with data where available (e.g., PipeLine). If we did not have accurate data for the beginning stock, we assumed no stock on hand for those cases. This assumption is justified because where inventory data were available, the quantities on hand indicated negligible stock levels on implants, especially in comparison with the growth of implant consumption that occurred in the following years. This assumption was also reasonable because our baseline year of 2010–2011 predates the price reduction agreements by several years (signed in 2013) and the other implants (Norplant and Sino-implant) did not represent a significant portion of procurements.

Estimates of Users of Implants
One estimate of the number of implants consumed was based on a combination of United

### TABLE 1. Source and Dates for Distribution-to-Client Data

<table>
<thead>
<tr>
<th>Country</th>
<th>Source of Direct-to-Client Distribution of Implants</th>
<th>Period Covering Direct-to-Client Distribution Data</th>
<th>Period Using Demographic Estimates Instead of Distribution-to-Client Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burkina Faso</td>
<td>Pipeline database</td>
<td>January 2010–June 2017</td>
<td>n/a</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>n/a</td>
<td>n/a</td>
<td>January 2011–December 2017</td>
</tr>
<tr>
<td>Pakistan</td>
<td>eLMIS</td>
<td>January 2010–December 2017</td>
<td>n/a</td>
</tr>
<tr>
<td>Uganda</td>
<td>n/a</td>
<td>n/a</td>
<td>January 2010–December 2017</td>
</tr>
<tr>
<td>Country A&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Annual Statistical Survey</td>
<td>September 2014–December 2017</td>
<td>January 2010–August 2014</td>
</tr>
<tr>
<td>Country B&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Pipeline database</td>
<td>January 2010–December 2017</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Abbreviations: DHIMS, district health information management system; DHIS 2, District Health Information System 2; eLMIS, electronic logistics management information system.

<sup>a</sup>Two countries have been anonymized to enable inclusion of their data in this article.
Nations population data, mCPRs, implant share of method mix or implant CPR, and couple year of protection (CYP) conversion factors. The mCPR, implant share of method mix, and implant CPR were collected for all women (with the exception of Pakistan, for which mCPR and method mix were available only for married women) from DHS and PMA2020 surveys and when used in the below formula can estimate users of implants. See the Appendix for the data and sources used for each country.

\[
\text{implant users} = \frac{\text{female population of reproductive age}}{\text{CYP}} \times \text{mCPR} \times \text{implant share of method mix}
\]

Or

\[
\text{implant users} = \frac{\text{female population of reproductive age}}{\text{CYP}} \times \text{implant CPR}
\]

Since our analysis reviewed products procured and available through public health and subsidized channels such as ministries of health and NGO and social marketing organization services, we did not include users who accessed implants from private retail markets. Our estimates for source mix of combined public and NGO sectors also came from DHS (or in the case of Pakistan from the electronic logistics management information system) and are included in the Appendix as well.

### Converting Users From Demographic Data to Implant Requirements

Multiple methodologies exist to estimate products required to serve a number of users of a particular method for a given period of time. We relied on the conventional method to convert the number of users of a product to quantities of products required for a period of a year used by Reality Check, a tool created by EngenderHealth and widely used by many family planning practitioners to project family planning trends (i.e., CPR and method mix) and the resources required to achieve them. To convert the number of users into estimated number of products required, we used the established concept of CYPs developed by the United States Agency for International Development and respective conversion factors. CYP factors are an estimate of protection provided by contraceptive methods during a 1-year period and yield an estimate of the average duration of use and contraceptive protection provided per unit of that method and therefore differ from the period of method efficacy. Although CYP factors were not intended to be used for procurement planning, they have been used in most methodologies to estimate product requirements from user data. Since all users would have required an implant at some time, the CYP is used as a proxy to determine the period of time in which they may have received the implant. Because hormonal implants are effective for multiyear periods, the CYP provided is per unit as opposed to units per CYP, as with shorter-acting methods. CYPs for longer-acting methods are also not as long as the full, potential duration of use of the method because they account for early discontinuation and other factors such as wastage. The CYP factors used in our model are presented in Table 3.

Since demographic survey results do not indicate brand share of implants, we used procurement, consumption mix, and/or forecast assumptions to estimate the share between the 2 brands. The final

### TABLE 2. Inventory Policy on Maximum Months of Stock Holdings for Each Level of the Supply Chain as Determined by Each Country’s Supply Chain Design

<table>
<thead>
<tr>
<th>Level</th>
<th>Burkina Faso</th>
<th>Ethiopia</th>
<th>Ghana</th>
<th>Kenya</th>
<th>Pakistan</th>
<th>Tanzania</th>
<th>Uganda</th>
<th>Country A</th>
<th>Country B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central medical stores</td>
<td>15 (combined)</td>
<td>5</td>
<td>12</td>
<td>30</td>
<td>14</td>
<td>9</td>
<td>6</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>Regional/hubs level</td>
<td>4</td>
<td>6</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>6</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>District level</td>
<td>5</td>
<td>4</td>
<td>n/a</td>
<td>n/a</td>
<td>3</td>
<td>6</td>
<td>n/a</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>SDPs</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>n/a</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Health posts</td>
<td>n/a</td>
<td>2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>System total</td>
<td>22</td>
<td>19</td>
<td>21</td>
<td>32</td>
<td>18</td>
<td>15</td>
<td>10</td>
<td>24</td>
<td>14</td>
</tr>
<tr>
<td>System total used in model</td>
<td>22</td>
<td>18</td>
<td>19</td>
<td>21</td>
<td>30</td>
<td>18</td>
<td>15</td>
<td>10</td>
<td>24</td>
</tr>
</tbody>
</table>

Abbreviations: n/a, not applicable; SDPs, service delivery points.
CYP factor used in the analysis applied country-specific estimated share of the 2 brands. For example, if a country procured, consumed, and/or forecasted a brand mix of 25% levonorgestrel and 75% etonogestrel implants, then the CYP used in the calculations used would be 2.825 as described in this calculation:

\[(25\% \times 3.8) + (75\% \times 2.5) = 2.825,\]

giving a CYP conversion factor of 1/2.825, or 0.354.

The Reality Check calculation to convert the number of users of a product to quantities of products required for a period of years is as follows:

\[(U_{Y2} - U_{Y1}) + (U_{Y1} \times CYP_{CF})\]

Where:

\(U_{Y1} = \text{Users in year 1}\)

\(U_{Y2} = \text{Users in year 2}\)

\(CYP_{CF} = \text{Couple year protection conversion factor} (\text{products per year of protection})\)

Our calculations required monthly figures; therefore, the above formula was adapted to yield monthly estimates:

\[(U_{M2} - U_{M1}) + \left(U_{M1} \times CYP_{CF} \times \frac{1\text{ year}}{12\text{ months}}\right)\]

Where:

\(U_{M1} = \text{Users in month 1}\)

\(U_{M2} = \text{Users in month 2}\)

Based on the adapted formula, a CYP of 2.5 for 1-rod, 3-year etonogestrel implant translates to 30 months of protection (with a CYP_{CF} of 1/30 or 0.3), and a CYP of 3.8 for the 2-rod, 5-year levonorgestrel implant translates to 45.6 months of protection (with a CYP_{CF} of 1/45.6, or 0.022). This monthly adaptation of the Reality Check calculations results in the same aggregated annual product total as the original Reality Check formula.

Final CYP factors used for each country have not been provided in this article because doing so would divulge confidential market information.

**Data on Distribution of Implants Directly to Clients**

We also collected data on direct-to-client distribution of implants when and where available. Direct-to-client distribution represents quantities of products actually dispensed/used. Typically, distribution of products to clients is tracked through records used at individual SDPs such as health management information systems (tracking the number of women who receive an implant insertion service) or logistics management information systems (tracking the dispensing/consumption of the implant product).

We chose to use actual dispensed-to-client data instead of demographic estimates, explained earlier, as an input to the model when available for answering the first and second research question. This input is described in Table 1.

**Country Supply Chain Parameters**

Country supply chain design and inventory management policies dictate the total system volumes required for the supply chain to respond effectively to variable demand. We reviewed country documents and spoke with country representatives to determine distribution points along the supply chain and respective maximum inventory stocking policies for the 9 countries included in the study. Inventory levels are typically measured in relative terms of demand—that is, months of stock rather than absolute quantities, given that the intent is to be able to assess how long the current stock levels would be expected to last, based on current and future demand. Table 2 documents the maximum months of stock inventory policies for each level of the supply chain in the 9 countries.

### Table 3. Couple Years of Protection and Conversion Factors

<table>
<thead>
<tr>
<th>Implant Brand</th>
<th>CYPs/Implant</th>
<th>CYP_{CF}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Etonogestrel, 1-rod (Implanon/Implanon NXT)</td>
<td>2.5</td>
<td>1/2.5 = 0.4</td>
</tr>
<tr>
<td>Levonorgestrel, 2-rod (Jadelle)</td>
<td>3.8</td>
<td>1/3.8 = 0.26</td>
</tr>
</tbody>
</table>

Abbreviations: CF, conversion factor; CYP, couple year of protection.
reviewed as well as the system total used in the models.

The total line reflects months of stock that are needed to ensure adequate system inventory levels at all levels of the supply chain, as determined by country supply chain design, and assumes that quantities procured are distributed to distribution points and SDPs as dictated by the supply chain design of the country. The total used in the model for Ethiopia and Kenya deviates from the system totals because not all products flow to the health posts in Ethiopia and maximum system inventory levels are not strictly adhered to in Kenya for procurement purposes. In Pakistan, although some SDPs receive products directly from the central level, those working in the system indicated that the majority of implants are being used in province(s) supplied from the district. Totals used were made in consultation with country implementers familiar with each system.

Although countries typically maintain their inventory levels between predetermined inventory parameters of minimum and maximum stock levels, our model used the system maximum as the benchmark for measuring the surplus or deficit.

All models start in January 2010 (with the exception of Ethiopia and Tanzania, which start in January 2011) and include procurements and estimates of product consumption information through the end of 2017 when data collection for the model was completed. Since we limited further assumptions of implant use past 2017, we used the results of the monthly difference between imputed and filled-to-max inventory up to December 2017, or last survey point available, minus the maximum months of stock required for system inventory. Table 4 represents the period of time used for the results of the analysis for answering research questions 1 and 2.

### RESULTS

#### Accuracy of Procurement Quantities

Three of the 9 countries, Ethiopia, Pakistan, and Country A, on average came closest to accurate procurement over the 5 to 6 years studied, while in 4 countries (Burkina Faso, Ghana, Kenya, and Tanzania), procurement volumes were on average lower than what was required to fill the supply chain to maximum inventory requirement levels. In Country B and Uganda, procurement volumes on average exceeded the need, as shown in Figure 5. Pakistan and Country B had very wide variations in average procurement quantities, while Burkina Faso, Kenya, Tanzania, and Country A had much narrower ranges of procurement quantity accuracy over the period of time reviewed. Figure 2 depicts the results of how accurately countries were able to procure implants through 4 to 6 years (referenced in Table 4) and compares maximum inventory needs of each country’s supply chain based on actual procurement to what should have been procured (filled-to-max inventory procurement quantity) for the estimated demand. Countries with asterisks in Figure 5 have actual dispensed-to-client data incorporated in the analysis. Any country that crosses the 0% line had periods of time when it had accurate stock, based on the analysis.

To interpret the results in Figure 5, 0% procurement error indicates that procurement quantities on average matched the need of the supply chain perfectly. Percentages below 0% indicate

<table>
<thead>
<tr>
<th>Country</th>
<th>Demographic and/or Dispensed to Client Data Collected</th>
<th>Monthly Model Results Used in Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burkina Faso</td>
<td>January 2010–December 2017</td>
<td>January 2010–August 2015</td>
</tr>
</tbody>
</table>

3 countries had accurate procurement, while 4 countries had volumes below needs and 2 countries had excessive volumes.
that procurement quantities were on average insufficient to fill the supply chain to maximum system inventory levels to meet the estimated demand. In contrast, percentages above 0% indicate that procurement quantities on average exceeded what was needed to fill the supply chain with system maximum inventory to meet the estimated demand. For example, in Tanzania procurement volumes for this period were 56% lower than what would have been required to maintain maximum inventory levels; however, this result may not mean that Tanzania was understocked, or had procurement quantities below what was required for maintaining minimum inventory levels, because this possibility was not included in this analysis. The results for Uganda were an anomaly. The data appeared to show on average a very high overstocking of implants. However, upon further review, all countries but Uganda have access to some level of dispensed-to-client data. These countries also use such consumption data to inform forecasting. Uganda, on the other hand, historically has relied heavily on demographic methodology for forecasting demand. A situation in which demographic estimates cannot be compared to dispensed or issued data may result in overstocking or the appearance of overstocking because we do not have accurate data on dispensed-to-client data of implants in Uganda.

The data also showed that inventory levels and related procurement volumes do not have a standard factor for relating inventory quantities to consumption rates across countries or even within a country through time. A situation in which demographic estimates cannot be compared to dispensed or issued data may result in overstocking or the appearance of overstocking because we do not have accurate data on dispensed-to-client data of implants in Uganda.

Had there been no actual dispensed-to-client data available, forcing us to use only lower demographic estimated implant consumption in the model (see results to research question 3), procurement quantity accuracy would actually appear to increase. The same procurement volumes would better fill a system to max levels when there appears to be a lower demand. Table 5 presents the comparison of these 2 results with all but one country showing a lesser degree of under-procurement or even over-procurement when using our demographic estimates in the model.

This increased likelihood of showing overprocurement when using demographic data for demand estimates may be why Uganda appears to be overstocked in Figure 5. Had there been dispensed-to-client data (to show higher consumption than demographic estimates), the system may in reality not have been overstocked.

**Standard Factor for Relating Consumption Data to Procurement Volumes**

Our results also demonstrated that there was no standard factor for relating inventory quantities to consumption rates across countries or even within a country through time as the program demand changes. Inventory needs in a month as a percentage of consumption in the same month varied from 18% to 1,356% and averaged about 180% for the countries and periods evaluated, as shown in Figure 6, almost double what was consumed. We believe that these results do not just pertain to implants only but to any product that does not have stable demand. Therefore, simply applying a percentage to demand to account for system inventory needs is likely to produce inaccurate procurement quantities.

The data also showed that inventory levels and related procurement volumes do not have a standard factor existed for relating inventory quantities to consumption rates across countries or even within a country through time.

**Simply applying a percentage to demand to account for system inventory needs is likely to produce inaccurate procurement quantities.**

---

**FIGURE 5. Procurement Quantity Accuracy Based on Maximum Quantities Required to Fill the Supply Chain**

Abbreviation: SD, standard deviation (average + SD is shown with the top horizontal line while average — SD is shown with the lower horizontal line).

Note: Uganda, which appears to have a very high overstocking of implants, is plotted on a separate, secondary y-axis to avoid obscuring the differences for the other countries.

*Dispensed-to-client data were incorporated in the analysis.
consistent relationship to consumption volumes. The range of percentages is explained by context-specific factors that influence the proportion of consumption/demand that should be in inventory (i.e., procured) and vary based on system design parameters and rates of growth or decline in consumption. The results also reflect the rapid rates of scale-up associated with the implant programs in countries. The country with the widest range, Pakistan, also had the highest percentage growth in implant users, increasing from 6 users in the first month of the analysis to 12,370 in a peak month.

Accuracy of Demographic Estimates
Comparison of demographic estimates of consumption to dispensed-to-client data for the 6 countries for which both sets were available revealed that demographic estimates (using the Reality Check methodology) were on average lower than the dispensed-to-client consumption based on data from all 6 countries. Comparisons were available for 435 months (over 36 years) of data. For each country, the average consumption of implant forecasts estimated using demographic estimates as a percentage of dispensed-to-client data were below

<table>
<thead>
<tr>
<th>Country</th>
<th>No Dispensed-to-Client Data Used in Model</th>
<th>Dispensed-to-Client Data Used, if Available, in Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burkina Faso</td>
<td>–29%</td>
<td>–83%</td>
</tr>
<tr>
<td>Ghana</td>
<td>–23%</td>
<td>–39%</td>
</tr>
<tr>
<td>Kenya</td>
<td>–111%</td>
<td>–112%</td>
</tr>
<tr>
<td>Tanzania</td>
<td>51%</td>
<td>–56%</td>
</tr>
<tr>
<td>Country A</td>
<td>–7%</td>
<td>55%</td>
</tr>
<tr>
<td>Country B</td>
<td>112%</td>
<td>–4%</td>
</tr>
</tbody>
</table>

a Bolded figures indicate the appearance of more accurate result, and negative figures represent under-procurement.

Abbreviation: SD, standard deviation.
100% and ranged from 47% to 96%, yielding forecast errors in using demographic estimates that ranged from 4% to 53%. The averages, range, and standard deviations were also calculated for each country and are provided in Figure 7.

To interpret Figure 7, in Tanzania, the estimated forecasts using demographic estimates on average represented only 48% of the number of implants dispensed to clients based on the electronic logistics management information system data reported at facilities; the average forecast error would have been 52% had forecasts been made using demographic estimates.

The findings for question 3, indicate that demographic estimates, using the Reality Check methodology, underestimate dispensed-to-client consumption of implants. Although this study did not evaluate the reasons for the discrepancy, studies evaluating the validity of using CYP conversion factors have highlighted challenges, particularly for condoms and long-acting contraceptives. This underestimate of implant demands using demographic estimates has implications for the use of this demographic methodology (including CYPs) for forecasting consumption for contraceptive implants.

Model Evaluation
To validate the results of the model, the outputs of the imputed ending inventory balance were compared to reported total system-wide full inventory in Burkina Faso, one country for which these data were available. Burkina Faso reports system inventory (central, regional, and SDP combined) in its PipeLine database and routinely updates the stock on hand at the time of physical inventory. The recorded system inventory data were available for the following periods: January 2010, January 2011, January 2012, July 2012, January 2013, July 2013, January 2014, July 2014, January 2015, July 2015, January 2016, and July 2016.

We compared these reported system inventories to the model output of imputed ending inventory for the same months and plotted the results in Figure 8. The resulting Pearson correlation coefficient of the 2 datasets is 0.97061, indicating a high degree of relationship between the 2 data sets.

Based on this analysis, we feel confident that the model outputs align with imputed stock-on-hand information on the ground. Similar quality data to validate model outputs were unavailable for other countries.

DISCUSSION
Our results show that the significant investments in procurement quantities for the rapidly growing implant programs in the 9 countries were justified based on consumption rates. The results also show that the investments were likely an important enabler for countries to achieve the high insertion rates for this new method and for implants to achieve a growing proportion of mCPR for the period reviewed. Further, the results suggest that most of the countries included in this evaluation

Demographic estimates underestimate dispensed-to-client consumption of implants.

The significant investments in procurement quantities for implants in the 9 countries were justified based on consumption rates.
historically have not been ordering quantities that meet their systems’ maximum inventory needs, given the length of their supply chains and the rate of growth in use of implants. However, countries that had estimated quantities of inventory below the system filled-to-max inventory levels should not be interpreted as having insufficient quantities or quantities below the system minimum stock levels because this research did not attempt to answer that question. The analysis also did not attempt to assess overstocking above maximum inventory levels within the in-country pipeline or by distribution level.

Study results also suggest that large stocks of product accumulating unused in national or sub-national warehouses do not explain the difference between quantities procured and insertion rates. Rather, the variation in the inventory-to-consumption ratio of 18% to 1,356% demonstrated in Figure 6 reflects the variation in efficiency of each country supply chain in moving contraceptives from ports of arrival through the in-country system to end users. With each additional level that the product has to travel through and with limitations in storage and distribution infrastructure, varying quantities of inventory are needed for each country to ensure the pipeline is filled at each level so that facilities can be resupplied regularly and clients can access an implant whenever they need one.

In addition, our results demonstrate that calculated consumption using demographic estimates (with the Reality Check methodology) as a percentage of dispensed-to-client data was below 100% in all 6 countries for which both types of data were available, and ranged from 47% to 96%. These forecast errors in using demographic data ranged from 4% to 53% and are likely one reason for underestimates, given that many programs use demographic estimates as the primary forecasting methodology.

Best practice recommends that countries use multiple methodologies of forecasting demand, review the strengths and weakness of each methodology, and compare results to select final forecasts outputs (either as a blended or soundest forecast).10,12 Yet some national programs may not be able to forecast using multiple methodologies and sources of data for 2 reasons. The first is the limited visibility and availability of consumption (dispensed to client and/or issues) and services data from routine information systems. When systems do not provide sufficiently complete or credible data, program managers may rely heavily on demographic data and estimates for forecasting. The second reason for not using multiple methodologies and sources of data is that the relevant skills and practices may not be widespread in many countries, even when data are available. Our results indicate that countries that rely solely on demographic data and use the demographic forecasting methodology (using the same methodology as in this analysis) may routinely under-forecast contraceptive implants and be unable to completely fill their supply chain with sufficient inventory.

It is important to consider the results of this overall analysis in the context of general supply chain management principles and as governments, donors, and partners consider investments in supply chain strengthening. As with any forecast, accuracy is never 100% and errors are typical. Inventory must be available in advance of

![Burkina Faso Scatter Plot of PipeLine Reported System Inventory and Imputed Ending Balance for the Same Reporting Months](image-url)
service to make product choice and increasing mCPR possible; in this scenario, if product demand is increasing, order quantities will continue to increase and the supply chain will not be “filled” until the growth in demand levels off to zero. This advanced time is equal to the length of the supply chain at maximum inventory requirements. Therefore, based on demand trends for implants, large quantities of inventory (and respective procurements) are needed to fill the supply chain and should not be considered excess stock.

Another relevant supply chain management concept to consider in relation to this study is cycle (or customer) service levels—the ability to meet client demand on the day of visit with the product desired. Service levels below 100% for any given period mean that not all demand was satisfied and there was a stock-out. Contraceptive stock-outs are one of the primary and recurring challenges that affect many countries and jeopardize women’s reliable access to family planning products. Stock-outs also pose a major risk to the ambitious goals of FP2020. While family planning programs have placed increased emphasis on ensuring no stock-outs, it is important to recognize that realistically, reaching 100% customer service level (or no stock-outs ever) is not achievable and requires quantities of inventory far in excess than resources allow.\(^{11,12}\) That said, ensuring that sufficient inventory is available (including sufficient safety stock at points of resupply) and supply chains can be responsive to any sudden changes in demand is the best way to minimize both the frequency and duration of stock-outs.

To decrease inventory and yet continue to provide the same or greater level of service, supply chains need to be more efficient and/or responsive. With changes to supply chain design, it is possible to increase inventory turns while still keeping customer service levels high and to reach goals. Several levers can be applied to achieve greater efficiency, including decreasing the number of distribution points or lead times between distribution points; increasing review and reorder frequency; holding stock at lower levels of the supply chain; and finally, reducing the uncertainty of lead times and demand variability.\(^{11}\) It is important to note that responsible redesign of supply chains requires understanding of all costs in the supply chain, not just inventory holding costs, because making changes without understanding all costs may increase total system costs.

The results also emphasize the importance of establishing and strengthening electronic logistics management systems that collect supply chain data and the use of these data for operational and management decisions—including for robust forecasting and quantification—to improve performance. Understanding true demand at the last link of the supply chain (e.g., at the SDP) is essential for managing supply throughout the chain. Without this information, supply chain managers have limited ability to increase performance and efficiency of their systems, which ultimately limits women’s access to family planning and their choice of method.

This study also raised a few considerations about the use of CYP when used with the Reality Check methodology to convert number of users to quantity of products, particularly for long-acting methods. CYP is widely used to estimate the number of couples protected from pregnancy based on the quantities of family planning contraceptives distributed, and it allows donors and programs to estimate the impact of the contraceptives provided.\(^{19}\) Although not developed for this purpose, the conversion factor has been widely used in demographic forecasting to convert number of users to products required, a methodology described earlier. Recent studies evaluating the validity of using the conversion factors to determine products required have highlighted challenges, particularly for condoms and long-acting contraceptives. This challenge likely reflects inadequate data on quantities typically used (as in the case of condoms) and duration of actual use (as in the case of long-acting reversible contraceptives).\(^{17,18}\)

Our results indicate that estimates using demographic data with the Reality Check methodology and CYPs set consumption for implants too low. Therefore, other methodologies, such as the Marie Stopes International Impact 2 model, should be tested to determine if they present better options for determining product needs from users.\(^{20}\)

### Assumptions and Limitations

The model uses a number of assumptions about the data and relationships. The system filled-to-max inventory procurement quantity assumes that the country has no resource limitations and that product is available at the top of the supply chain (procured) to flow through the system at the right time. The model also assumes that products flow through the in-country system and are supplied downstream to distribution points and SDPs as the system design dictates.

Another assumption is that forecasts and supply plans are developed solely at the central level

---

**Responsible redesign of supply chains requires understanding of all costs in the supply chain.**
and are accurate. Assuming forecast error of any percentage, a norm with all forecasts and supply chains, would require additional quantities of inventory to account for this variance in demand. Further still, if each level of the system was doing its own forecast to determine its demand, additional inventory would be needed at each point of the supply chain to account for the further uncertainty in demand at each level.

Although directly measuring dispensed-to-client data quality was not possible, we reviewed data received to ensure there were no obvious data entry or other errors. We had access to data from other countries in addition to those included, but used peer judgment of quality to only include countries that provided better quality dispensed-to-user data in the analysis. We also assume that the supply chains are as described in Table 2 and that countries are following supply chain designs as described. Lastly, our model assumes that clients consume all products. The model does not account for wastage and/or products used for training, and it assumes no loss or damage of products, which would need to be removed from the system and deducted from stock levels before reaching a client.

CONCLUSION

This research should provide assurances that rapid increases in procurement quantities of implants for countries where we have data have not resulted in system overstocks to date. Our results should also reinforce the idea that the relationship between procurement quantities and consumption levels cannot be accurately assessed without understanding the country supply chain, inventory control parameters, and current and future demand. Further, the need to procure large quantities of products well in advance of when they are needed may be reasonable, although it does raise questions about opportunities to redesign supply chains such that inventory (and related holding costs) can be reduced and what the cost and service level benefits of a redesign might be.

This study also provides ideas for future research. First, a future research study could explore other methodologies for demographic estimates and the validity of using CYP factors when converting users to number of products used, particularly for long-acting methods such as contraceptive implants. Second, future research could evaluate countries with longer, multilevel supply chains to cost out the supply chain current state, including the cost of holding inventory, and potential alternative scenarios to determine if supply chain redesign (reducing levels, increasing order frequency) could help reduce the inventory required to meet demand while also minimizing total systems costs.

Acknowledgments: We would like to thank the inSupply staff who conducted the initial supply chain data collection and analysis. We would also like to thank Maryjane Lacoste, of the Bill & Melinda Gates Foundation, and Deborah Dull, of GE Digital, for their support of this research. Finally, we are grateful to the staff at the Ministries of Health in Burkina Faso, Ethiopia, Ghana, Kenya, Tanzania, Pakistan, and Uganda who shared and provided the necessary information required to complete the research represented. Without the availability and visibility of the supply chain information, this research would not have been possible.

Funding: Bill & Melinda Gates Foundation

Competing Interests: None declared.

REFERENCES


Peer Reviewed
Received: January 4, 2019; Accepted: April 17, 2019


© Akhlaghi 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-19-00017
### APPENDIX. Demographic Data Used in Estimation of Implant Users

<table>
<thead>
<tr>
<th>Country</th>
<th>Data Source</th>
<th>mCPR</th>
<th>Implant CPR</th>
<th>Implant Share of Method Mix</th>
<th>Public-Sector and NGO Source Mix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burkina Faso</td>
<td>DHS 2010</td>
<td>14.3</td>
<td>2.9</td>
<td></td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>PMA-R1 2014</td>
<td>15.7</td>
<td>43.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMA-R2 2015</td>
<td>18.6</td>
<td>34.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMA-R3 2016</td>
<td>21.5</td>
<td>41.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMA-R4 2017</td>
<td>21.9</td>
<td>43.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethiopia</td>
<td>DHS 2011</td>
<td>18.7</td>
<td>2.3</td>
<td></td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>PMA-R1 2014</td>
<td>22.5</td>
<td>15.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMA-R2 2014</td>
<td>23.8</td>
<td>20.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMA-R3 2015</td>
<td>25.6</td>
<td>20.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMA-R4 2016</td>
<td>26.5</td>
<td>22.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DHS 2016</td>
<td>24.9</td>
<td>5.7</td>
<td></td>
<td>96.8</td>
</tr>
<tr>
<td>Ghana</td>
<td>DHS 2008</td>
<td>13.5</td>
<td>0.7</td>
<td></td>
<td>85.4</td>
</tr>
<tr>
<td></td>
<td>PMA-R1 2013</td>
<td>14.3</td>
<td>13.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMA-R2 2014</td>
<td>14.6</td>
<td>14.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DHS 2014</td>
<td>18.2</td>
<td>3.7</td>
<td></td>
<td>94.1</td>
</tr>
<tr>
<td></td>
<td>PMA-R3 2014</td>
<td>18.1</td>
<td>12.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMA-R4 2015</td>
<td>23.4</td>
<td>14.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMA-R5 2016</td>
<td>21.7</td>
<td>18.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMA-R6 2016</td>
<td>21.3</td>
<td>23.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kenya</td>
<td>DHS 2008</td>
<td>28</td>
<td>1.3</td>
<td></td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>DHS 2014</td>
<td>39.1</td>
<td>7.1</td>
<td></td>
<td>82.1</td>
</tr>
<tr>
<td></td>
<td>PMA-R1 2014</td>
<td>42.5</td>
<td>18.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMA-R2 2014</td>
<td>40.3</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMA-R3v2015</td>
<td>46.5</td>
<td>21.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMA-R4 2016</td>
<td>46</td>
<td>23.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMA-R5 2016</td>
<td>44.2</td>
<td>27.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMA-R6 2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pakistan</td>
<td>DHS 2006-2007</td>
<td>21.7</td>
<td>0.1</td>
<td></td>
<td>eLMIS data by month</td>
</tr>
<tr>
<td></td>
<td>DHS 2012-2013</td>
<td>26.1</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tanzania</td>
<td>DHS 2010</td>
<td>23.6</td>
<td>1.8</td>
<td></td>
<td>94.8</td>
</tr>
<tr>
<td></td>
<td>DHS 2015-2016</td>
<td>27.1</td>
<td>5.6</td>
<td></td>
<td>97.8</td>
</tr>
</tbody>
</table>
## APPENDIX (continued).

<table>
<thead>
<tr>
<th>Country</th>
<th>Data Source</th>
<th>mCPR</th>
<th>Implant CPR</th>
<th>Implant Share of Method Mix</th>
<th>Public-Sector and NGO Source Mix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uganda</td>
<td>DHS 2011</td>
<td>20.7</td>
<td>1.9</td>
<td></td>
<td>87.5</td>
</tr>
<tr>
<td></td>
<td>PMA-R1 2014</td>
<td>21</td>
<td></td>
<td>11.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMA-R2 2015</td>
<td>26.2</td>
<td></td>
<td>12.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMA-R3 2015</td>
<td>25.9</td>
<td></td>
<td>14.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMA-R4 2016</td>
<td>27.5</td>
<td></td>
<td>13.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DHS 2016</td>
<td>27.3</td>
<td>4.7</td>
<td></td>
<td>84</td>
</tr>
<tr>
<td>Country A</td>
<td>DHS 2006</td>
<td>4.5</td>
<td></td>
<td>n/a (100% used)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DHS 2012</td>
<td>11</td>
<td>0.3</td>
<td>n/a (100% used)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMA-R2 2016</td>
<td>12.6</td>
<td></td>
<td>17.6</td>
<td></td>
</tr>
<tr>
<td>Country B</td>
<td>DHS 2007-2008</td>
<td>16.3</td>
<td>0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DHS 2010</td>
<td>25.2</td>
<td>3.6</td>
<td>94.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DHS 2014-15</td>
<td>27.8</td>
<td>4.7</td>
<td>92</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CPR, contraceptive prevalence rate; DHS, Demographic and Health Survey; eLMIS, electronic logistics management information system; mCPR, modern contraceptive prevalence rate; n/a, not applicable; PMA, Performance Monitoring and Accountability.
The Challenges of Transition From Donor-Funded Programs: Results From a Theory-Driven Multi-Country Comparative Case Study of Programs in Eastern Europe and Central Asia Supported by the Global Fund

George Gotsadze, a Ivdity Chikovani, a Lela Sulaberidze, a Tamar Gotsadze, a Ketevan Goguadze, a Nertila Tavanxhib

ABSTRACT

Introduction: In the era of declining development assistance for health, transitioning externally funded programs to governments becomes a priority for donors. However, the process requires a careful approach not only to preserve the public health gains that have already been achieved but also to expand on them. In the Eastern Europe and Central Asia region, countries are expected to graduate from support from the Global Fund to Fight AIDS, Tuberculosis and Malaria in or before 2025. We aim to describe transition risks and identify possible means to address them.

Methods: Using a theory-based conceptual framework—Transition Preparedness Assessment of Tuberculosis and HIV/AIDS programs—we investigated transition-related challenges through a health systems lens in 10 countries of the Eastern Europe and Central Asia region during 2015–2017. Study findings were derived from systematic collection of quantitative data on socioeconomic indicators and disease epidemics as well as qualitative data from in-depth interviews with 264 stakeholders. These findings were then compared with other donor transition experiences documented elsewhere.

Results: We found numerous common transition challenges, such as poor monitoring of a country’s macroeconomic performance along with weakness in estimating financial needs for successful transition; limited political will of governments to replace donor-funded programs; punitive legislation criminalizing certain behaviors and constraining the government’s ability to allocate funds and contract civil society organizations essential to providing services for key populations; limited coordination function of governments and weak decision-making power of coordinating mechanisms obscuring the latter’s future role; and inadequate function of national procurement and supply chain management systems undermining an uninterrupted supply of quality-assured drugs and commodities. These challenges are compounded by the risks related to health workforce management leading to specialist shortages and/or inadequately skilled and qualified professionals and by limited funding for critical surveillance activities.

Conclusion: The complex and multidimensional transition process requires a multipronged approach through well-planned collective and coordinated responses from global, bilateral, and national partners in coming years. Other similar transition processes may provide guidance. Although no “one-size-fits-all” approach exists, previous experiences highlight a need for both early planning and monitoring of the transition along several key dimensions. Issues that could threaten the maintenance of health gains include ongoing stigma against key populations; continued heavy reliance on external funding in some countries, especially for preventive services; the institutional viability of the country coordinating mechanisms; and emerging difficulties with procurement of quality drugs at reasonable prices.

INTRODUCTION

During past decades, development assistance for health (DAH) increased substantially and fostered progress toward global health goals.1 Increased investments in health have helped countries to improve maternal, newborn, and child health, and to reduce the spread of HIV/AIDS, malaria, tuberculosis (TB), and other major infectious diseases.2-3 These investments were channeled through bilateral and multilateral agencies, as well as through global health initiatives such as Gavi, The Vaccine Alliance; the Global Fund to Fight AIDS, Tuberculosis and Malaria (Global Fund); and so forth.

a Curatio International Foundation, Tbilisi, Georgia.
b UNAIDS, Geneva, Switzerland.

Correspondence to Lela Sulaberidze (l.sulaberidze@curatio.com).
In the field of HIV/AIDS, TB, and malaria, the Global Fund, among others, was pivotal in achieving public health gains by investing close to US $38.7 billion or approximately 9% of DAH during 2002–2016. These investments helped to enhance the national coordinating structures in charge of national responses to TB, HIV/AIDS, and malaria; advance public-sector capacity; mobilize civil society and community organizations and engage them in service delivery especially for the most at-risk and vulnerable populations; and expand and scale up preventive, diagnostic, curative, and supportive interventions. Above all, these investments raised public awareness about epidemics and promoted approaches based on human rights.

Following the 2008 global financial crisis, DAH levels stagnated, and declining trends have been observed since 2013. This decline has triggered debates within the donor community about the gradual transition of donor-funded programs to country ownership. These discussions have also affected the Global Fund, which led the Executive Director to raise the following concern with the Board:

With some humility, we can admit that in development work, including global health, there have been a lot of exits but not many successful transitions. Programmatic and financial sustainability takes time, planning and a balanced portfolio of trades and investments along the development continuum.

In other words, the public health gains achieved by recipient countries seem to be at risk unless the transition from donor support in general, and from the Global Fund in particular, is well planned and executed. This issue takes on even greater importance in the context of the Sustainable Development Goals, which include commitments to universal health coverage12 and renewed commitments to Alma Ata.13,14

Consequently, we decided to examine the potential transition challenges in countries in Eastern Europe and Central Asia (EECA) that are expected to graduate from Global Fund support in or before 2025. Using theory-based comparative case studies from 10 countries (Armenia, Belarus, Bulgaria, Georgia, Kosovo, Kyrgyzstan, Moldova, Turkmenistan, Ukraine, and Uzbekistan), we investigated the programmatic areas within a broader country context that could be at greatest risk during transition. We hope these findings will facilitate discussions on potential solutions going forward as well as inform discussions around transition within the Global Fund, among donors and recipient countries.

The Challenges of EECA

Since 2000, EECA has made significant progress addressing the challenges posed by a growing epidemic of TB and HIV/AIDS. However, the threats remain and the region requires even greater attention as countries head toward transitioning from donor support. In this section, we briefly describe the most significant epidemiological and other trends that need attention, highlighting the importance of the region from an epidemiological perspective.

Although the rate of new HIV infections is decreasing globally, it more than doubled in EECA between 2006 and 2015. Due to the low level of testing coverage, almost a third of the people infected are not aware of their HIV status. The HIV epidemic is concentrated predominantly among key populations (KPs) that are driving the growth of the epidemic—primarily, people who inject drugs followed by men having sex with men (MSM). While support from donors, especially the Global Fund, has led to significant progress in developing, delivering, and scaling up preventive, diagnostic, curative, and support services for KPs, the coverage rates of HIV prevention programs within the region are still low. According to United Nations Office on Drugs and Crime, almost a quarter of the people injecting drugs around the world reside in the EECA region, or approximately 2.9 million people. However, coverage with opioid substitution therapy remains below 5% in all but 3 states, and access to needle and syringe programs, while variable across countries, remains below the recommended 200 clean needles and syringes per person who injects drugs per year.

Despite the efforts of the past decade to scale up treatment coverage, only 21% of people living with HIV/AIDS (PLHIV) in EECA were receiving treatment in 2015, which is far below the global average of 53%. Thus, the rate of new infections continues to outpace antiretroviral therapy enrollment and undermine the goal to end AIDS as a public health threat by 2030.

HIV prevalence data for MSM are variable and grossly misleading for the region due to weak surveillance systems. The reported prevalence among MSM in some countries is as high as 20.7% in Georgia and as low as 0.8% in Armenia. High HIV rates are compounded by variable rates of self-reported condom use.
ranging from 49% in Moldova to 81.6% in Kyrgyzstan. The percentage of MSM reporting using a condom the last time they had anal sex with a male partner varies from 80.4% in Armenia to 61.2% in Moldova, with low usage facilitating the spread of infection.

Fighting societal stigma and implementing HIV response are impeded by conservative legislation and political and cultural barriers related to same-sex relationships, drug use, and sex work. These challenges often drive both behavior and services underground, reducing the scale and impact of disease programs. Stigma and discrimination also continue to hinder access to HIV prevention, treatment, and care services for KPs, thereby exacerbating social inequalities.

The situation with TB is similarly poor. Even though the broader European region (including EECA) has had declining TB rates since 2000, which has led to a reduced TB burden, multidrug-resistant TB (MDR-TB) has emerged as a substantial public health threat. Nine of 30 countries with the highest MDR-TB burden in the world are in EECA, representing about 20% of the global MDR-TB burden (approximately 350,000 individuals). The proportion of MDR-TB cases among new and previously treated TB cases in the region is significantly above the global average, with 19% in new and 55% in previously treated cases, compared with 4.1% and 19%, respectively, as of 2016. Despite universal treatment coverage for TB and MDR-TB, the treatment success rate in the region is below regional and global targets, which indicates a need to improve treatment program performance.

High rates of TB and HIV co-infection plague the region. Furthermore, TB was the most common AIDS-defining illness in the EECA region in 2015, and the number of incident TB cases co-infected with HIV almost doubled (from 5.5% to 9%) between 2011 and 2015. TB remains a significant cause of death among PLHIV: the rate of TB-related deaths among PLHIV increased by 3.6% annually between 2011 and 2015. Action is required. Civil society as well as communities in EECA are appealing to the West to pay adequate attention to these mounting threats as donors consider transitioning the region away from external support.

Managing donor transitions responsibly is crucial to ensure that the public health gains attained with donor support are not only preserved but also expanded, which is essential to adequately deal with the TB and HIV threats in the region.

### Methods

Following Bennett et al., we define transition as the formal handing over of a donor-funded health program to one or more local partners in a way that ensures critical elements of the program are sustained over time. Given how central the notion of transition is to sustainability and the long-term effects of the entire development enterprise, surprisingly little literature and relatively limited empirical evidence exist with regard to what constitutes good transition practice and what needs to be done during the process. Transition is increasingly seen as a process rather than an outcome because the programs evolve through complex adaptive systems, while responding to changing contexts; as a result, the activities implemented during transition need to meet ever-changing needs. To achieve sustainability after donor funding has ended and to retain or expand the public health gains that were achieved during donor support, the transition process must be adequately understood and managed. Although much of the research in the field of transition is retrospective, we decided to undertake an evaluation of the transition issues inside a group of 10 countries before transition based on our case studies arising from short-term consultancy work. We examined TB and HIV/AIDS programs supported by the Global Fund, with the exception of the HIV/AIDS program in Turkmenistan, and our evaluation of 19 programs in 10 countries has informed the findings of the current study.

Curatio International Foundation, with the financial support from the Global Fund, developed a theory-based conceptual framework for a transition preparedness assessment (TPA) of Global Fund-supported programs (Figure). The TPA framework builds on existing sustainability frameworks, including the U.S. President’s Emergency Plan for AIDS Relief (PEPFAR) Sustainability Index and Dashboard, using a health systems lens. The TPA framework looks at external and internal environments, their subdomains, and components such as economic and political support, financial and human resources, information systems, governance, accountability, service delivery, organizational capacity and state of transition planning (the TPA framework components are described in the Figure). Using quantitative and qualitative information/indicators, the TPA framework helps to identify the sustainability-related issues that require attention during transition. The framework was operationalized as a tool that applies a scoring
system for each indicator under a component. Each component is assigned a risk category for transition (high, moderate, or low) based on the measurement of its indicators. A more detailed description of the TPA tool methodology is available elsewhere.38 The data for this article were collected during the period 2015–2017 from 10 countries in the EECA region using the TPA framework and tool. Although the TPA tool has also been used in Jamaica, Morocco, and the Philippines, we maintain a regional focus for this paper on EECA, specifically Armenia, Belarus, Bulgaria, Georgia, Kosovo, Kyrgyzstan, Moldova, Turkmenistan, Ukraine, and Uzbekistan.

To apply the TPA tool, we used a mixed-methods approach consisting of desk review, quantitative data collection from public databases, and qualitative data collection through in-depth interviews. The desk review focused on legislative and policy documents, program performance, evaluation/reviews, program expenditure reports, and other types of reports. Quantitative data on socioeconomic indicators and disease epidemics were collected from databases belonging to the World Bank, the Joint United Nations Programme on HIV and AIDS (UNAIDS), and the World Health Organization (WHO). In-depth interviews (using standard guidance/tools adjusted to specific country contexts) were conducted with key stakeholders who were selected through snowball sampling including government officials; the principal recipients of Global Fund grants; national HIV and TB programs representatives; members of civil society organizations (CSOs), international NGOs, and UN agencies; and representatives of the Global Fund secretariat in Geneva. A total of 264 respondents were interviewed (see Supplement Table 1 indicating the number of country representatives). To ensure robustness of the data, the country findings were triangulated across the different data sources, where possible. Finally, the TPA framework helped to systematically extract and compare data/information across the country case studies to identify the common program areas exposed to transition risks.

By revealing the more generalizable trends across countries, we intend to (1) describe the areas at greatest risk during the transition, and (2) contribute to global knowledge about the most expected transition challenges in order to enrich the debates around transition planning by donors/funders and countries.

**RESULTS**

In this section, we organize the findings around the external environment, and then present the results concerning the internal/program...
environment with its subdomains and components. The Table provides a summary matrix of the most common barriers of transition found across the studied countries in the EECA region. The barriers are structured by TPA components and categorized by high, moderate, or low risk for transition. More granular information by country is provided in Supplement Table 2.

Economic and Political Environment

The economy determines a country’s capacity for national budgets, including funding for the health sector and disease programs. The economies of the countries studied performed well during 2010–2016, although the average annual per capita GDP growth rate varied from 0.47% (Ukraine) to 7.8% (Turkmenistan). During the same period, government revenues (not including grants) averaged 28.9% of GDP across the sample with significant variability: the lowest proportion (23.1% of GDP) occurred in Armenia, while the highest (35% of GDP) was in Ukraine. Besides the variable macroeconomic performance, the political will of governments to fund health in general and disease-specific programs in particular is a prerequisite to ensure that they take responsibility for replacing donor funding during the transition. The importance of a nation’s health in public financing is highly variable when considered as the proportion of the state budget devoted to health. This proportion ranged between 5% in Georgia and 13.5% in Belarus in our sample (based on averages for the 2010–2014 period). At the same time, the share of public spending in total health expenditure ranged from 18.7% in Georgia to 71.5% in Belarus (average figures for 2010–2014). The study also found that the estimations of future financial needs for the national programs were either of poor quality and/or not available.

Laws and regulations criminalizing specific behaviors are another important component of the external environment because they could pose significant challenges to the transition process. All countries in this study, except Bulgaria, criminalize the behaviors associated with KPs to varying degrees. In some countries such as Armenia, Belarus, and Moldova, illicit drug consumption is a criminal offense. In Belarus, changes in the legislation have created additional barriers and reinforced stigma and discrimination, such as mandatory registration of all drug users in an electronic database accessed by law enforcement body, obligatory disclosure of HIV status when a person seeks medical care, and compulsory HIV testing. In Uzbekistan, homosexual sex is still criminalized and sex work is subject to an administrative fine.

Financing of Disease Programs

The dependence of the programs on donors is sizable but also variable across countries. For example, the HIV/AIDS program in Moldova was 68% externally funded for 2015, while in Uzbekistan, the HIV/AIDS program received only 29% from donors during the same year. We found that the TB program had less dependence on external funding, ranging from 15% in Turkmenistan and in Uzbekistan (2015) to 51% in Kyrgyzstan (2016). Dependency on external support becomes even more apparent when the disease program components are considered. For example, KPs are prioritized in almost all national HIV/AIDS strategic plans of the sample countries, but harm reduction and other prevention activities largely rely on external (Global Fund) financial support. Furthermore, public spending on KPs over past years did not grow, and if any increases occurred, they were marginal.

In all cases except Moldova, low threshold services are not funded by public sources; Belarus, Moldova, and Ukraine provide little cofinancing from the national budget for opioid substitution therapy services, and only Georgia and Bulgaria fund these services predominantly from the national budget. Moreover, all other countries entirely depend on the Global Fund for funding opioid substitution therapy service provision. In sum, all countries exhibit significant dependence of preventive interventions from external sources. Similarly, TB programs in 10 countries depended heavily on Global Fund support for drugs especially for second-line drugs and diagnostics (Supplement Table 2). By 2016, close to 40% of TB allocations from the Global Fund went for pharmaceutical products, another 10% for nonpharmaceutical products, and 5% for equipment. However, TB programs planned to reduce external funding by increasing national budget spending starting from 2017 but the actual results still have to be validated.

Social Contracting

Continuing engagement of NGOs/CSOs emerged as one of the critical impediments in the transition process in all countries. During the past decade, Global Fund and donor funding to NGOs/CSOs helped to increase their role in delivering critical preventive, care, and support services, primarily
### TABLE. Summary Matrix of Transition Barriers Across the Study Countries

<table>
<thead>
<tr>
<th>TPA Component</th>
<th>Major Barriers</th>
<th>Country(^a) and Year of Assessment</th>
<th>ARM</th>
<th>BLR</th>
<th>BGR</th>
<th>GEO</th>
<th>KOS</th>
<th>KGZ</th>
<th>MDA</th>
<th>TKM</th>
<th>UKR</th>
<th>UZB</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>External Environment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economic environment</td>
<td>Poor macroeconomic performance with reduced potential for budget revenues constrains the government’s ability to allocate more funds to health and disease programs. If this is compounded by limited political will to increase funding for health services from the state budget, the risk to transition increases.</td>
<td>HIV/TB</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Political environment</td>
<td>Laws, regulations, or policies criminalizing certain behaviors (in full or in part) constrain governments’ ability to allocate budgetary resources towards services needed for KPs. Punitive legislation constrains the government’s abilities (as well as political will) to allocate budget funds and contract NGOs/CSOs for services focused on KPs involved in the criminalized activities. The societal stigma arising from “traditional values” shapes homophobic attitudes in wider society and significantly influences the development and enforcement of national legislation and constrains the government’s ability to allocate resources for KPs.</td>
<td>HIV</td>
<td>H</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td><strong>Internal Environment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inputs</td>
<td>High dependence on external financing for important HIV preventive services targeting KPs. Proper estimates of the future financial needs of disease programs are usually lacking or are of poor quality.</td>
<td>HIV</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>Human resources</td>
<td>Weak human resource planning and development practices. Health workforce shortage in TB services due to staff aging and poor replacement. Weak HR capacity for surveillance and disease program management. Lack of well-functioning educational programs supporting necessary human resource production. Inadequate integration of donor-supported training programs in education system.</td>
<td>HIV</td>
<td>M</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>H</td>
</tr>
<tr>
<td>Health information systems</td>
<td>High dependence of HIV second-generation surveillance and population size estimation studies on donor financial support. Inadequately advanced surveillance systems coupled with limited analytical capacity.</td>
<td>HIV</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td><strong>Governance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Governance (coordination mechanism)</td>
<td>Limited legal basis of CCMs within sovereign legislation posing risks to coordinated, multisectorial responses to disease epidemics.</td>
<td>HIV/TB</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td>M</td>
<td>H</td>
<td>H</td>
<td>M</td>
</tr>
<tr>
<td>Program</td>
<td>Lack of public financing rules and regulations governing CSO/NGO contracting in the health sector.</td>
<td>HIV/TB</td>
<td>M</td>
<td>H</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td>M</td>
<td>H</td>
<td>H</td>
<td>H</td>
</tr>
</tbody>
</table>

\(^a\) ARM = Armenia; BLR = Belarus; BGR = Bulgaria; GEO = Georgia; KOS = Kosovo; KGZ = Kyrgyzstan; MDA = Moldova; TKM = Turkmenistan; UKR = Ukraine; UZB = Uzbekistan.
to KPs. Their role increased more prominently in the HIV programs than in TB, for which active CSO engagement is relatively recent but growing. Most NGO/CSO-delivered services are externally funded and face the threat of reductions in Global Fund support. Such reductions would decrease the volumes of delivered services by NGOs/CSOs unless state budgets pick up the tab and/or greater efficiency in service delivery is achieved.

According to Aceso Global, social contracting is the process by which government resources are used to fund entities that are not part of government (e.g., CSOs) to provide services. Social contracting may have various names and slightly different mechanisms between countries. Regardless of the terminology used, social contracting mechanisms typically involve a legally binding contract, in which the government agrees to pay a CSO for services rendered and the CSO agrees to provide certain deliverables in exchange.43

Deeper examination of the context for social contracting revealed that all the countries in our sample have enacted legislation that allows government contracting of NGOs/CSOs. Furthermore, all the countries in our sample practice civil society contracting in sectors other than health. However, within our sample, only Armenia, Georgia, Kosovo, and Moldova use social contracting in the health sector. Belarus introduced changes to the laws allowing for social contracting of NGOs for prevention of HIV/AIDS and other communicable diseases only in mid-2017, while other countries have not yet tried NGO contracting. Thus, while NGO/CSO contracting frameworks are included in the national legislation, in most instances the mechanisms either are not relevant for procuring preventive and other services in the health sector and/or they do not specify the necessary details for the allocation and/or disbursement of public funds. Examples include the lack of specific national standards or guidelines defining services currently being delivered by NGOs/CSOs; the lack of methodology (and capacity) necessary for estimating the budget requirements for the services and/or for evaluating the adequacy of quoted prices in the bids; the lack of clear tendering procedures, bid selection criteria, and contracting terms and conditions (including programmatic and financial reporting requirements for NGOs/CSOs); and the lack of guidelines/procedures for monitoring the volume and quality of delivered services. In addition, few CSO/NGOs in the countries have sufficient financial and technical capacity to engage in public procurement processes.

**TABLE.** Continued

<table>
<thead>
<tr>
<th>TPA Component</th>
<th>Major Barriers</th>
<th>Country and Year of Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Limited CSO/NGO capacity to engage in the public procurement processes.</td>
<td>Armenia (ARM) 2017</td>
</tr>
<tr>
<td></td>
<td>Lack of proper national registration for GFATM supplied drugs under the grant.</td>
<td>Belarus (BLR) 2015</td>
</tr>
<tr>
<td></td>
<td>Weak national procurement systems with inadequate capacity for forecasting supply needs for the programs.</td>
<td>Bulgaria (BGR) 2015</td>
</tr>
<tr>
<td></td>
<td>Absent or weak national quality assurance (system for drugs and weak postmarketing surveillance, monitoring of adverse drug reactions.</td>
<td>Georgia (GEO) 2015</td>
</tr>
<tr>
<td></td>
<td>Structural impediments of pharmaceutical market leading to higher prices.</td>
<td>Kyrgyzstan (KGZ) 2017</td>
</tr>
<tr>
<td></td>
<td>HIV/TB</td>
<td>MDA 2016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Turkmenistan (TKM) 2017</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ukraine (UKR) 2015</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Uzbekistan (UZB) 2017</td>
</tr>
</tbody>
</table>

Abbreviations: ARM, Armenia; BLR, Belarus; BGR, Bulgaria; CCM, country coordinating mechanisms; CSO, civil society organization; GEO, Georgia; GFATM, Global Fund to Fight AIDS, Tuberculosis, and Malaria; HR, human resources; KGZ, Kyrgyzstan; KOS, Kosovo; KP, key population; MDA, Moldova; TB, tuberculosis; TKM, Turkmenistan; TPA, transition preparedness assessment; UKR, Ukraine; UZB, Uzbekistan.

43 The red color denotes high risk for transition or existence of many or significant barriers, yellow denotes moderate risk for transition or existence of few barriers, and green indicates a low risk for transition or an absence of major barriers.

4 For country-level findings, the risk categories are indicated by letters: H for high risk, M for moderate risk, and L for low risk for transition. A blank cell indicates that the topic was not studied.

All countries studied have legislation allowing government contracting of NGOs/CSOs, but few use such contracting in the health sector.
Governance and Coordination Function
Without strong governance and cross-sectoral coordination arrangements in place, managing the complex process of transition will be difficult. A lack of vision for continuing intersectoral coordination that ensures civil society engagement during and after transition emerged as an area of concern in our review. While almost all stakeholders highlighted the importance of country coordinating mechanisms (CCMs), the feasibility of maintaining CCMs in the post-Global Fund era is questionable due to the currently weak institutional placement of CCMs in sovereign governance structures. For example, a CCM is placed under the central government only in Belarus, Bulgaria, and Uzbekistan. In other countries, CCMs are created by government resolutions but are largely placed under the ministries of health, and other sectors typically do not attend the meetings. Furthermore, most respondents questioned the ability of CCMs to make decisions and ensure their implementation because CCM decision-making powers are not aligned with the sovereign governance systems/rules. These concerns were further exacerbated by the CCMs’ almost complete dependence on Global Fund funding, without a clear vision for how they would be supported financially in the post-Global Fund era.

Human Resources
Weak human resource planning and development practices along with a shortage of TB specialists due to staff aging and poor replacement are challenges that pose risks to transition if not addressed adequately. Almost all the studied countries are experiencing active labor migration; have failed to plan for an adequate number of health professionals; lack deployment and staff motivation policies; and have health workforce shortages affecting provision of TB care. In all the countries, at least 20% of TB specialists are in the prereirement and retirement age groups. In Georgia, the proportion is 30%, while it is even higher in Kyrgyzstan (44%). Due to low financial motivation and professional risk exposure, younger people throughout the region are not willing to pursue this profession. Additionally, HIV and TB training courses that were developed with the support of development partners are not yet fully integrated into national undergraduate, postgraduate, or continuing professional development programs. The exception is Moldova, where all levels of medical education have been updated according to the latest guidelines. Some countries have sporadically updated continuing medical education curricula, while undergraduate and residency programs lag behind. The countries lack a repository of all necessary training materials and qualified master trainers. If these issues are not addressed during transition, once external funding ends, the medical education system will produce health professionals without adequate professional knowledge and skills.

Procurement and Supply Chain Management
Considering transition-related procurement and supply chain management issues under the “organizational capacity” component of the TPA framework (Figure), we found that an uninterrupted supply of quality-assured drugs and diagnostics seems to be at risk during and after the transition. Countries in the sample still rely on Global Fund procurement mechanisms such as voluntary pooled procurement—a Global Fund strategic initiative that aggregates order volumes on behalf of participating grant recipients to negotiate prices and delivery conditions with manufacturers, mainly for second-line antiretroviral (ARV) drugs. For TB drugs and commodities, they also rely on the global drug facility (GDF), a procurement mechanism that is the largest global provider of quality-assured TB medicines, diagnostics, and laboratory supplies to the public sector. Some countries, such as Belarus, Kyrgyzstan, and Armenia, have already established alternative supply channels using public financing, albeit with emerging challenges. Specifically, purchasing drugs and commodities through alternative and/or national suppliers, have resulted in the following:

- **Prices higher by 80% or more for ARV and TB drugs and diagnostics.** Belarus paid higher prices for ARV drugs; depending on the medicine, the prices were 2 to 8 times those of the Global Fund. In Uzbekistan, regional HIV centers purchased test systems through a decentralized, local procurement process at prices that were 70%–80% higher than those of the Global Fund. Ukraine paid a higher price for ARV drugs prior to amending legislation in 2015. These developments are not unique to HIV and TB or to the EECA region; they were also observed in other countries that had graduated from the Global Fund or Gavi, when price increases were significant for purchases through local/regional suppliers. Price increases naturally demand higher amounts of public funds and thus emerge as a threat to transition and sustainability.
The ability of health information systems to capture the comprehensive data necessary for planning, nationwide coverage, and practice varies across countries.

Transition processes are complex and multidimensional, encompassing numerous domains as well as actors within and outside the health system.

- **Questionable quality of the commodities supplied, especially drugs, which could negatively affect treatment outcomes and lead to drug resistance.** When dependent on Global Fund grants, and therefore on the commodities supplied through quality-assured international systems, countries almost always used one-time import waivers to procure drugs as opposed to obtaining proper market authorization following national legislation. Consequently, quality-assured drugs supplied through Global Fund systems were not registered in the graduating markets, and when countries moved to public financing and applied sovereign procurement rules, they were frequently left with only local suppliers that may not have had quality-assured products. This challenge is compounded by weak or absent national quality assurance (QA) systems for drugs, coupled with weak postmarketing surveillance and monitoring of adverse drug reactions. While all EECA countries have QA requirements for pharmaceuticals, usually based on international recommendations, local stakeholders question the degree of successful implementation and the efficiency of these systems.

- **An interrupted supply of drugs.** Numerous factors adversely affected drug supply, including poor planning and quantification of the needed commodities (most often due to a lack of capacity); delays in delivery from local suppliers winning national tenders; administrative-bureaucratic challenges arising from national public procurement and financing rules; and little or no interest from pharmaceutical companies to participate in public tenders due to the countries’ small market size or other reasons, such as corrupt practices.

Thus, national procurement of drugs and commodities require specialized market knowledge, institutions, and skills that still need to be built or enhanced in some graduating countries if the quality of supplied drugs and commodities is to be secured during and after the transition.

**Health Information Systems**

Health information systems that produce critical epidemiological and program data are pivotal for adequate national response planning and management. External assistance throughout the past decade has been critical in building and developing these systems for evidence-based decision making. Most countries have moved towards electronic TB information systems, supported by external assistance, but HIV information systems lag behind. However, the ability of these systems to capture the comprehensive data necessary for evidence-based planning, nationwide coverage, and application in practice varies across countries.

Repeated waves of bio-behavioral surveillance (BBS) studies among KPs, and more recently population size estimation (PSE) studies have been critical for measuring national HIV/AIDS program outcomes. They also offer important information for advocacy and program management. Although the methodologies have variable rigor, these studies are vital for tracking disease prevalence and population behavior, especially among KPs. Countries in totality depend on external financial support to implement BBS and PSE studies.

To conclude, the findings reveal that Global Fund program transitions are expected to face multiple risks, and therefore a more carefully planned approach to transition is required.

## DISCUSSION

The evaluation of transition processes in our sampled countries revealed that transition processes are complex and multidimensional, encompassing numerous domains as well as actors at all levels within and outside the health system. In this section, we discuss the differences and similarities of major transition challenges and relate the lessons learned from the countries of the EECA region and countries that have already transitioned from Global Fund support. We also discuss countries’ experiences of graduating from the support of other bilateral and multilateral donors, including the United States Agency for International Development (USAID), the Bill & Melinda Gates Foundation, and Gavi. Such comparisons help to identify potential solutions that have worked in other parts of the world and could be instrumental for the Global Fund going forward.

The available literature suggests that a collaborative and coordinated process between donors and countries is important in planning for transition. This approach helps to generate political commitment from the government and stronger buy-in from national stakeholders who help plan and manage successful transition by alleviating or mitigating potential negative consequences. The process includes 4 essential...
elements which are described in greater detail later in this section:

- Early planning with the government to reach a mutually acceptable and time-bound transition plan divided into phases with clear milestones, which allows for sufficient time for any adjustments and corrections necessary to proactively mitigate existing or emerging risks.
- Aligning donor-funded program components with government structures and funding modalities before transition.
- Building government capacity through active technical assistance and management support, and budgeting for adequate support during and after the transition.
- Developing and using a framework for monitoring the transition process along with a mechanism for ensuring mutual accountability between a donor and a country.

A transition plan represents a necessary instrument for securing a government’s political commitment to use its funds and capabilities and gradually replace donor-funded services, commodities, and management responsibilities during and after graduation. Our findings corroborate findings in the literature that indicate achieving successful transition requires close monitoring of a country’s macroeconomic performance and its fiscal space, along with the current and expected health sector and disease spending needs. We also found that accurate and reliable estimates for the future financial needs of disease programs are usually lacking or are of poor quality, which will likely negatively affect adequate financial planning for transition. Therefore, estimating national financial needs during and after donor graduation, while simultaneously monitoring a country’s fiscal space, seems necessary when making transition decisions and/or planning for the duration of the transition process. The literature suggests that transition planning requires forecasts of expected program cost at least 5–10 years into the future in order to account for realistic and not purely aspirational programmatic goals.48

However, political commitment should not only be financial; it also has to include an obligation to enact legislative and regulatory changes to address the barriers. Apparently, most diagnostic and treatment services (except second-line drugs for TB patients) are already funded and delivered by the governments. However, preventive services are highly dependent on external support.

In addition, we found that punitive legislation constrains governments’ ability and political will to allocate budget funds and contract NGOs/CSOs for services focused on KPs. It is important to recognize that “traditional values” influence the development and enforcement of punitive legislation that makes KPs more marginalized and vulnerable to HIV. Unless such legislation and regulations are amended during the transition process, achieving sustainable handover of certain preventive services would be very difficult. Also, punitive legislation imposes access barriers to services, unless services are delivered by CSOs in a friendly environment for KPs. Examples include delivering needle exchange services where illicit drug use is criminalized; offering services to sex workers and MSM, including condom distribution, where criminal or administrative liability exists for these individuals; and requiring mandatory parental consent for adolescents for HIV testing.

Most attempts to build political commitment must be targeted at national governments because the budgets, laws, policies, and regulations that can sustain a health program in the long term often flow from governments and are closely interlinked.49 We found that CCMs in the studied countries lack adequate placement in the government hierarchy, which leads to weak coordination and decision-making power. In addition, their future role after transition is vague. The experiences of countries that have transitioned from Global Fund support suggest that CCM and their coordination function fades after transition and limits CSO involvement in the decision-making process.50 We speculate that if CCM funding was shifted to the Ministry of Health budget, it would still be uncertain whether the secretariats could maintain their independence and cross-sectoral coordination roles. Consequently, the Global Fund practice of seeking CCM-approved transition plans, in a context in which the legal powers of current CCMs and their abilities after transition are uncertain, may require rethinking.

Transition plans found to be central to sustainability have included the following elements: strategic prioritization of critical program areas; budgeted recommendations for action; and a clear timeline and phases for graduation, with associated benchmarks and indicators to assess progress.10,51 Planning for transition proved useful for the Avahan program in India,7 Gavi graduation,8 and USAID-supported family planning programs.6,10 However, developing a transition plan is both a technical and a political process that...
requires adequate time for development and proportional investments. Such a plan requires transition risk analysis and thorough planning; fiscal space analysis and adequate estimation of the required resources through different budget scenarios, agreement for equitable sharing of the financial burden between donors and the country in question during the transition; and planning for the integration of disease programs with national funding streams. Most importantly, it also requires a process of intensive dialogue and the establishment of transparent accountability mechanisms between domestic and international funding organizations.

Before graduation from donor support, a progressive reduction of dependency on external support is needed along with the alignment of donor-funded program components with government structures and funding modalities. Based on experiences from other programs, these steps may require coordinated program harmonization with existing services during the transition. This harmonization can include adapting program services and implementing arrangements for planning, management, financial reporting, and monitoring and evaluation, to foster integration with the national program or host environment. To facilitate program transition, technical, managerial, and cost elements of programs need to be aligned with government norms. As noted in the 10 countries studied, Global Fund–supported programs operated their own supply line of quality-assured drugs and diagnostics, which were imported to the country with one-off waivers instead of following national regulations and securing proper market authorization. Critical surveillance activities such as BBS and PSE studies were solely donor funded instead of being integrated as a routine component of national program surveillance supported by the government; and CSO/NGOs to deliver critical services to KPs were funded from external sources. Challenges in so-called “social contracting” were prominent in our sample and are also well described elsewhere in the literature.

Funding NGOs/CSOs from a state budget is challenging on several counts: (1) legal and societal barriers limit governments (and their political will) in allocating and spending budget funds on services, primarily focused on KPs; (2) public financing mechanisms/regulations allowing governments to contract NGOs/CSOs using budget funds in the health sector are lacking; and (3) NGOs/CSOs have a limited capacity to engage in government procurement processes to manage public funds due to national regulations. Thus, alignments during transition are multipronged and require durable mechanisms (legislative, procurement systems, or otherwise) to be established for the continuous supply of quality-assured commodities during and after transition with national funding. The process must include establishing CSO/NGO contracting systems/mechanisms/regulations to ensure uninterrupted service delivery to KPs after Global Fund graduation. Further, surveillance activities currently funded by the Global Fund must be fully integrated and operated by the national program and funded through the national budget.

Experiences from other countries show that when external support is withdrawn, surveillance data production and dissemination deteriorate. A striking example is Croatia, which hosts a WHO collaborative center for HIV/AIDS surveillance; however, after the Global Fund transition, HIV/AIDS reporting became irregular, leaving the international community with a knowledge gap on HIV within KPs. Thus during the transition, it becomes essential not only to shift funding onto the government, but also to ensure that critical information is still being generated and broadly disseminated.

The multiplicity of transition risks illustrated in this study highlights the need for further enhancement of national capacity through the provision of technical assistance and management support. The shift of health program responsibility from donor to program recipient means that the capacity previously supplied by donors must be replaced and/or adapted in line with the priorities and capacities of local actors. Therefore, as transition progresses it will require a steady reduction of investments in commodities to redirect the focus of donor resources to technical assistance and support, which should be adequately budgeted and funded during and after the transition. Based on our findings, technical support areas could range from supporting legislative amendments and enhancing advocacy efforts to improving management and stewardship capacity and procurement and supply management systems. The needs for technical assistance will vary from country to country, although some common areas will emerge, such as CSO/NGO contracting or amending national legislation/regulations to facilitate the uninterrupted supply of quality-assured drugs and commodities.

Finally, monitoring transition is essential; it must start before the transition begins and should follow the entire process. The approach to monitoring might differ depending on the purpose it
To manage transitions responsibly, the Global Fund needs to fully and strategically exploit its funding and partnership model at the global, regional, and local levels.

First, it is essential for the Global Fund Board to fully understand transition risks and allow for a gradual transition process over time that is linked not only to GDP per capita and disease burden but also to other important and measurable indicators. Such indicators obviously need to be elaborated and reflected during the sustainability transition and in cofinancing, and probably in the eligibility policy and potentially in the allocation formula as well. However, the appetite of the Global Fund Board to consider other elements beyond GDP and disease burden or to revise allocation formula is not reflected in its decisions or documents.

Second, the Global Fund has to develop its capability to negotiate a transition with the country government, and not only with the CCM, and to enforce the negotiated agreements. For the time being, Global Fund policies do not reveal a readiness to explore alternative mechanisms for legal engagement with sovereign governments instead of CCMs.

Third, the Global Fund’s approach to transition does not seem to be sufficiently nuanced to ensure gradual and smooth transitions. Numerous concerns exist across the range of stakeholders, which indicates a need for closer attention and eventual reflection on the part of the Global Fund about the complexities related to transition in the Board’s established policies and/or in guidance notes issued by the Global Fund secretariat.

Fourth, since the Global Fund is a partnership model without a country presence, it needs to strategically exploit the competencies and advantages of its partners to provide tailored technical assistance and support to countries. For some transition challenges, the support could be delivered at a regional level (for similar problems across the country group) by engaging a regional partner through a partnership arrangement. Examples include involving WHO to help countries improve national drug legislation in a way that facilitates the supply of quality drugs and commodities; enhancing procurement planning capabilities through regionally delivered training workshops and/or tools; and applying a regional grant-making mechanism involving capable regional watchdog organizations that use a regional platform for country-level operations. The organization would then be able to hold national governments accountable for their promises in the transition plan or involve advocates to push...
for specific changes in the policy and/or practice also spelled in the transition plan.

Finally, the Global Fund could strategically exploit the capabilities of partners that have a more significant presence or ability to deliver assistance at the country level, such as the German BACKUP initiative, France’s 7%, PEPFAR, USAID, and others. In close cooperation with such a partner, the Global Fund could identify country-specific areas for support and seek their assistance in helping country governments implement the required changes.

To conclude, only concerted efforts on the part of the Global Fund and its partners, coupled with well-planned transition and coordinated support to the countries involved, can ensure a smooth and sustainable transition without undermining the public health gains achieved with donor help. However, this transition cannot be achieved unless the Global Fund Board and especially its donor constituency fully recognize the complexities of transition and demonstrate readiness to take the bold steps necessary to revise transition policy and procedures.

Acknowledgments: Support for this manuscript was provided by the Curatio International Foundation. The authors would like to acknowledge the support, cooperation, and guidance provided by Sara Bennett, Michael Borowitz, and Nicolas Cantau who shared their unique insights on transition for different programs that helped inform the TPA framework and tools. Ana Amaya contributed to the development of the TPA framework and to pilot testing in 4 countries. Further, we would like to extend our thanks to the individuals who worked on country case studies: Ketevan Chkhatarashvili, Mzia Tabatadze, Tamar Gabunia, and Tina Zardiasvili. We also thank all interviewees who kindly provided insight into their experiences during transition.

Funding: The authors report financial support from the Global Fund and UNAIDS while conducting of the country case studies. Preparation of the article was financially supported by Curatio International Foundation.

Competing Interests: None declared.

REFERENCES


Peer Reviewed
Received: October 29, 2018; Accepted: April 18, 2019


© Gotsadze 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-19-00042
Efficacy of a Digital Health Tool on Contraceptive Ideation and Use in Nigeria: Results of a Cluster-Randomized Control Trial

Stella Babalola,a Caitlin Loehr,b Olamide Oyebuia, Akinsewa Akiode,c Allison Mobleyd

A mobile digital health tool piloted in Kaduna City, Nigeria, was efficacious in promoting positive contraceptive attitudes and encouraging women to adopt a modern contraceptive method, thus showing potential for reducing unmet need in Nigeria.

ABSTRACT

Background: Contraceptive prevalence in Nigeria remains among the lowest in the world, which substantially contributes to the country’s high maternal and child mortality. Mobile phone technology penetration has increased considerably in Nigeria, opening opportunities for programs to use this medium for reaching their intended audience with health-protective information.

Methods: In 2017, the Health Communication Capacity Collaborative conducted a cluster-randomized control trial in Kaduna City to assess the efficacy of the digital health tool Smart Client on ideational and behavioral variables related to family planning. Twelve wards in the city were randomly assigned to intervention (6 wards) and control (6 wards) arms of the study. A total of 565 women aged 18–35 years were randomly selected from study wards and consented to participate in the study. At recruitment, the women completed a baseline survey. The women in the intervention group were registered to receive 1 welcome call, 13 program calls, and 3 quiz calls on their mobile phones. Each of the program calls had several segments, including introduction, drama episode, and friend-to-friend chat. The last quiz call included evaluation questions. Women in the control arm received no intervention. The efficacy of the intervention was assessed using both per-protocol and intent-to-treat differences-in-differences techniques.

Results: The intervention and control arms were equivalent in terms of key sociodemographic characteristics, with the exception of religion. Attrition was a major challenge in the study. On average, participants receiving the intervention listened to 7.2 drama episodes but only 2.6 personal stories and 1.1 sample dialogues. The results of both per-protocol and intent-to-treat analyses show that the intervention was efficacious in improving relevant ideational and behavioral outcomes. For example, the intent-to-treat results show that the intervention increased women’s perceived level of confidence to discuss family planning with a provider by 27.7 percentage points and modern contraceptive prevalence by 14.8 percentage points.

Conclusion: This efficacy assessment showed that using an interactive voice response-based digital tool that includes drama is a viable option for promoting positive ideation about family planning and increasing contraceptive use in Nigeria. Significant lessons learned from this efficacy trial include informing participants at the time of recruitment of what the opening segment of the calls will sound like to avoid the calls being mistaken for telemarketing calls and intensive testing prior to scale-up to avoid potential attrition due to technical issues.

BACKGROUND

Despite decades of government and donor investments in family planning service provision and demand generation, the contraceptive prevalence rate in Nigeria remains one of the lowest in the world. Findings from the 2016–2017 Multiple Indicator Cluster Survey revealed that only 13.4% of in-union women of reproductive age were using any contraceptive method, while only 10.8% reported using a modern method.1 Compared with data from the 1999, 2003, and 2013 Demographic and Health Surveys, these latest survey results indicate that contraceptive prevalence in Nigeria has not increased meaningfully over the last 20 years.2−3 Huge differences among the states and regions have also persisted. For example, in 2016, modern contraceptive prevalence varied between 0.2% in Yobe state in the northeast and 30.0% in Oyo state in the southwest.1 Low contraceptive use is a key contributor to the high levels of fertility, maternal mortality, and infant mortality observed in Nigeria.4−5 Indeed, contraception has long been recognized as a crucial intervention for improving maternal and child health outcomes and economic...
opportunities for women. Use of contraceptive methods has also been shown to positively influence gender-related dynamics and economic empowerment for women.

Lack of access is a key barrier to contraceptive use among Nigerian women. Studies have identified several other demand- and supply-side factors that affect contraceptive use in Nigeria and elsewhere in Africa. Literature abounds on the role of supply-side factors including service availability, distance to services, cost of services, and provider bias, interpersonal skills, and technical competence. On the demand side, sociodemographic characteristics of the women (e.g., age, parity education, religion), household and conjugal variables (e.g., partner’s characteristics, household wealth, type of marriage, woman’s autonomy), and community norms have also been shown to affect contraceptive use. The extant literature contains evidence of the important role of ideational (psychosocial) variables in the decision to use contraceptives. The ideational variables that have been found to be associated with contraceptive use include contraceptive awareness, attitudes, spousal communication, perceived self-efficacy to take actions related to contraceptive use, fear of side effects, misconceptions about contraceptives, perceived social support for contraception, and family size ideals.

For men and women, adopting a contraceptive method is often the result of a process that involves discussions and decision making about whether to practice family planning, what method to use, where to obtain it, and whether to continue using it. Throughout this process, a person’s fertility desires, perceptions about contraceptives, efficacy beliefs, and perceived social support for family planning may affect the decisions taken. The person may seek out information on family planning, talk with their partner, and discuss experiences with family and friends. At some point in this process, the person is likely to visit with a family planning provider.

Communication is a core skill running throughout this decision-making process—communicating with one’s partner, family and friends, and a health care provider. In most parts of Nigeria, however, communicating about sex, fertility desires, and use of family planning methods is not culturally appropriate. Furthermore, women and men often lack the skills they need to communicate effectively about such personal and sensitive subjects. Family planning demand generation programs are often designed to address the ideational, normative, and health systems-related factors that hinder women and men from seeking family planning services and adopting contraception. The programs rarely focus on preparing clients to be active and engaged communicators during their interactions with the family planning service provider. In many countries and settings, efforts made to improve providers’ communication skills and provide client-centered counseling have led to some improvement in client engagement, but the client is still often dependent on the provider to lead this process. As a result, female clients are often passive participants in family planning counseling, resulting in discussion and decision making being led by the provider.

The Health Communication Capacity Collaborative (HC3) was a 5-year project implemented between 2012 and 2017 by the Johns Hopkins Center for Communication Programs with funding from the United States Agency for International Development (USAID). One of the focal areas of HC3 was strengthening the capacity of its partners in developing countries to implement effective social and behavior change communication programs to improve contraceptive uptake. In this regard, the HC3 family planning team developed a digital tool, Smart Client or Beta Life in Nigeria, for implementing partners to use to increase the number of family planning clients who are informed, empowered, and confident. These “smart clients” are expected to be able to engage with providers and talk about their family planning needs. The Smart Client tool is inspired by earlier works on family planning clients’ coaching. The global proliferation of mobile technologies has led to their successful use for enhancing women’s knowledge about their health and increasing demand for various health services in many countries, including Bangladesh, Cambodia, Ghana, Kenya, Pakistan, South Africa, and Tanzania. Indeed, mobile phone-based interventions have been documented to promote behaviors related to child immunization, prevention of mother-to-child transmission of HIV, newborn health, pregnancy and infant care, and family planning. With digital health interventions, it is possible to reach a large segment of the intended audience with targeted health information. Moreover, digital tools can offer a more personal and private experience, which is particularly important in accessing information on a sensitive subject such as family planning. A woman may be hesitant to listen to a radio or television program because others might hear it as well, or she may not attend a community event for similar reasons. However, mobile phone-based approaches
can offer users a private space to access information. The advantages of digital health are particularly relevant in low-resource settings and for people with limited access to the traditional sources of health information (radio, television, community events), including women and other marginalized groups. Mobile phone penetration, both in terms of active subscriptions and unique subscribers, has increased conspicuously in Nigeria in recent years: from 23% of the population being unique subscribers in 2010 to 49% in 2017.48,49 This high level of penetration makes the use of this technology potentially effective for disseminating health information across various audience groups in the country.

The HC3 family planning team leveraged the power of basic mobile technology to develop a digital health tool to prepare women to become smart clients and encourage them to talk with their provider and partner about contraceptive methods. Here we report on the findings of the pilot test of the tool in Kaduna City, Nigeria. The analyses assess the effects of exposure to the digital tool on contraceptive ideation and use among women of reproductive age.

THE INTERVENTION

The Smart Client digital health tool was designed to inform, empower, and promote smart clients by reaching them directly through mobile phones. The tool is based upon Social Learning Theory, which posits that people learn from each other through observation, imitation, and modeling. The Smart Client tool therefore uses fictional role models, who demonstrate the desired behaviors and behavior change process in a drama format, as well as personal stories and examples of smart client dialogues. This approach allows the intended audience to observe an action, understand its consequences, and become motivated to repeat and adopt it. While drama is a common approach used in behavior change communication, it is usually delivered via television, radio, or community theater. This digital health tool explored how drama could be adapted to basic mobile phones via interactive voice response (IVR), using shorter and simpler storylines in a series of episodes while maintaining the fictional serial drama style. IVR was chosen as the delivery channel because it is accessible to audiences regardless of the type of mobile phone they have (e.g., smartphone or basic phone) and irrespective of their level of literacy. Given the lack of evidence for delivering entertainment-education content via IVR with the aim of promoting behavior change, the HC3 project tested and evaluated this approach.

The Smart Client digital health tool was designed to be delivered via mobile phone and included 17 prerecorded calls: 1 welcome call, 13 regular program calls, and 3 quiz calls interspersed. The order of the calls is outlined in the Box. Each regular program call includes 5 segments:

1. Brief welcome in which the host characters provide an introduction to the short drama included in the call
2. A short drama that follows a couple, Laila and Musa, and some of their friends as they face challenges and make decisions about contraceptive use
3. A “friend-to-friend chat” in which the hosts reinforce the key messages included in the drama segment and ask the user a quiz question
4. An optional personal story segment that requires the listener to use the keyboard to indicate whether they are interested in listening. Personal stories focus on diverse experiences related to family planning that correspond to the key message of the short drama episode.
5. Sample dialogue is another optional segment that features a friendly provider and a client, modeling what to expect during a family planning clinic visit and how to discuss needs, preferences, and concerns.

During the 3 short quiz calls, participants were asked a few brief questions to reinforce key messages, evaluate user understanding of content, and encourage user engagement. In addition, users received a short message service (SMS) reminder about the key message from each call.

The tool and its content were pretested in a focus group discussion setting among a group of women representative of the intended audience in Kaduna. Pretest participants shared feedback on the content—whether it was realistic, acceptable, and relevant—as well as their overall perceptions about the tool. The feedback served as the basis for finalizing the tool and its content prior to production and rollout. For the efficacy testing reported in this manuscript, the tool was given the name Beta Life, the content of the digital health tool was recorded in Hausa, and all SMS messages were written in Hausa.

The IVR platform was programmed so that users were preregistered and calls would be pushed to them on a schedule (every day, every other day, or twice per week) and time of day...
BOX. Description of the Smart Client Program

Call 1: Welcome call. The users listen to an introduction about the tool explaining how it works and what to expect from the content. They answer 3 questions regarding age, frequency, and time to receive calls using their numeric keypad. The pre-intervention questions were posed at the end of this call and study participants used their numeric keypad to respond.

Calls 2–7: Regular calls. These are part of the 13 regular calls. Users listen to the first 3 segments and then can choose to listen to any of the optional segments:

1. Brief welcome and introduction to the story by friendly host characters, a female and male.
2. Short episode of the serial drama, which follows a cast of characters including a couple, Laila and Musa, along with their family and friends, who all face different situations and challenges related to using family planning methods.
3. “Friend-to-friend” chats, in which the host “friends” deliver follow-up messages and tips related to the core message and the drama and ask the user a quiz question. Some messages in this segment are tailored for male and female users, based on their user preferences set on enrollment, or tailored to the user response to the question.
4. Personal story. This segment is optional, requiring users to “press 1” to hear the content. Personal stories, told by females and males, express diverse experiences with family planning that correspond to the key message of the episode.
5. Sample dialogue. This segment is also optional, requiring users to “press 2” to hear the content. Sample dialogues feature a friendly provider and a client or a couple, modeling what to expect during a visit to a family planning clinic and how to discuss needs, preferences, and concerns.
6. Wrap up and quiz question. To conclude the call, the hosts repeat key messages, ask users 1 to 2 question(s) to evaluate understanding of key messages, and sign off with a reminder to listen to the next call in the series.

Call 8: Short quiz call. Participants listen to the host ask 4 questions and answer these questions using their numeric keypad.

Calls 9–11: Regular calls. These are part of the 13 regular calls. Users listen to the first 3 segments and then can choose to listen to any of the optional segments as described above.

Call 12: Short quiz call. Users listen to the host ask up to 5 questions and answer using their numeric keypad.

Calls 13–16: Regular calls. These are part of the 13 regular calls. Users listen to first 3 segments and then can choose to listen to any of the optional segments as described above.

Call 17: Short quiz call. Users listen to the host ask 6 questions and answer using their numeric keypad. The post-intervention questions were asked at the end of this call and study participants used their numeric keypad to respond.

(morning, afternoon, or evening) of their choice. If the call came at an inconvenient time or if the user wanted to relisten to a call, they had the option of “flashing” or initiating a dropped call to the platform phone number and they would be immediately called back and could listen to the call that they missed or most recently listened to. For the purposes of the pilot test, users were not able to listen to calls out of sequence.

METHODS

Study Design and Data

A cluster-randomized control trial was used for this study. Clusters (wards of residence) were randomly assigned to one of two intervention conditions: receive the digital intervention or receive nothing. The required sample size was determined based on the proportion of women who had discussed contraceptive use with their spouse in the last 12 months. Since the value of this indicator was unknown in the study population, we assumed it to be 50% in our calculations since this level provided maximum variability. We also assumed that this indicator would increase by 15 percentage points among the women in the intervention group, and the required sample size was therefore 240 women for each arm. Estimating a loss to follow-up rate of 20%, we deemed it necessary to recruit 300 women into each arm. This number would provide a 90% power to detect a difference of 15 percentage points between the intervention and the control groups in the proportion of women who had discussed family planning with their husband or partner.

To recruit women into the study, we randomly selected 6 wards from each of the 2 local government areas (LGAs) in Kaduna metropolis—Kaduna North and Kaduna South. The study wards have comparable access to family planning services. Three wards from each local government area were randomly assigned to the intervention group and 3 to the control group. Trained female field agents, fluent in Hausa, went door to door in sample wards to identify eligible women, explain the purpose and method of the study, obtain informed consent, and recruit participants.

Consenting participants in the intervention group were registered to receive the Smart Client calls. The pre-intervention survey for the intervention group was administered as an automated survey at the end of the first call. The post-
intervention survey for this group was also an automated survey that directly followed the last call, between 3 and 11 weeks after they started the intervention depending on the frequency of the calls. The control arm did not receive the Smart Client intervention but received 2 calls on their mobile phone: one at the beginning of the study with the automated pre-intervention survey and the other 6 weeks later with the automated post-intervention survey.

At the time of recruitment, after informed consent was obtained in person, each participant completed a pre-study questionnaire to provide information on her age, number of children, religion, marital status, LGA and ward of residence, address, preferred nickname to be used during the study, primary and secondary cell phone numbers, and whether she shares a phone with anyone. Data from the pre-intervention and post-intervention survey calls, as well as user analytics collected by the IVR platform, were combined with pre-study data to conduct our analyses.

Setting and Participants
The study took place in North and South Kaduna LGAs of Kaduna State, Nigeria, from March 7, 2017, to June 5, 2017. The 2 LGAs are urban and make up the Kaduna metropolis. Residents included a mixture of Muslims and Christians, although the residents of Kaduna North are predominantly Muslim, while Kaduna South is predominantly Christian. Kaduna metropolis had an estimated total of 1.3 million inhabitants in 2017 and is a melting pot for various Nigerian ethnic groups. While the predominant ethnic group in the city is Hausa, the metropolis also includes large proportions of Yoruba, Igbo, Fulani, Gbaju, and other Nigerian ethnic groups. Secondary analysis performed by the lead author of survey data collected by Measurement Learning & Evaluation, in 2015 revealed that the majority (78.6%) of the women in the city had postprimary education while one-fifth had tertiary education. In the same survey, 21.0% of women of reproductive age reported using a modern contraceptive method, while 6.5% reported using a traditional method.

The intended audience for the digital tool is women of reproductive age. As such, women eligible for recruitment into the study were those with the following characteristics: aged between 18 and 35 years and not currently using a nonbarrier contraceptive method (e.g., pill, intrauterine device, implant, emergency contraceptives, tubal ligation, vasectomy, Lactational Amenorrhea Method), owned a mobile phone or had access to one, resident in Kaduna City, and fluent in Hausa.

Variables
The ideational and behavioral outcomes assessed in this manuscript include the following:

1. Considerations for desired family size—defined as having ever given thought to the number of children desired
2. Perceive self-efficacy for communicating with a family planning provider—defined as reporting a high level of confidence in one’s ability to discuss one’s concerns about contraceptives with a provider
3. Spousal communication about family size—defined as discussion of desired family size with one’s spouse in last 6 months
4. Spousal communication about contraceptive methods—defined as discussion of contraceptive methods with one’s spouse in last 6 months
5. Misinformation rejection—defined as rejection of the misconception that contraceptives can harm the womb
6. Current modern contraceptive use—defined as currently using any modern contraceptive method

Statistical Analysis
Data from the automated surveys (e.g., pre-intervention and post-intervention) and user analytics (e.g., number of calls received, number of episodes and segments completed) were combined with the demographic information collected at the time of recruitment and analyzed using summary statistics to compare ideational and behavioral outcomes among participants in the intervention or control groups. To assess the short-term effects of the digital health tool, difference-in-differences (DID) analytic method was employed. Note that each relevant outcome is measured in both the intervention and the control groups at 2 points in time: at the beginning and at the end of the study. DID evaluates the significance of the difference in gains over time between the intervention and control groups. More formally, the DID model is as follows:

\[ \delta = (Y_{1p} - Y_{1c}) - (Y_{0p} - Y_{0c}) \]

where \( \delta \) is the difference-in-difference estimator; \( Y_{1p} \) is the relevant outcome at the end of the study
for the intervention group; \( Y_{0p} \) is the relevant outcome at the beginning of the study for the intervention group; \( Y_{1p} \) is the relevant outcome at the end of the study for the control group; and \( Y_{0c} \) is the relevant outcome at the beginning of the study for the control group.

To strengthen the claim about the causal effect of the tool on assessed outcomes, the analyses controlled for relevant sociodemographic variables in the estimation of DID. Specifically, the estimation models controlled for the following variables: religion, current age, parity, education, marital status, and number of days elapsed between the pre-study interview and the end of the study interview.

The findings reported in this manuscript were derived from both per-protocol and intention-to-treat DID. In the per-protocol DID, only participants that met eligibility criteria were recruited into the study and only those that completed the post-study assessment were included in the analysis. The choice to do per-protocol analysis was due to the high level of attrition and because of the heterogeneity between the women who participated in the post-study survey and their peers that were lost to follow-up. Nonetheless, in conformity with CONSORT (Consolidated Standards of Reporting Trials) recommendations\(^{31} \), we also performed intention-to-treat analysis with all the women recruited into the study and who participated in the baseline survey. In the intention-to-treat analysis, eligible and recruited participants are included in the analysis, irrespective of whether they completed the post-study. The outcome was not measured for women who did not participate in the endline survey. For the intention-to-treat analyses, at post-study, baseline responses to ideational and behavioral questions were attributed to the women lost to follow-up since these responses were the most recent and only outcome information that we had for them. The significance of the intention-to-treat analysis should strengthen the claim about the efficacy of the intervention.

Furthermore, for the intervention group, user analytics were analyzed to track usage patterns (e.g., number of calls, average length of time listened to segments, navigation patterns, number of questions answered in quizzes, number of episodes heard) and gather general feedback on the user experience with the tool.

**Ethical Considerations**

The study was approved by the Johns Hopkins Bloomberg School of Public Health Institutional Review Board and by the National Health Research Ethics Committee in Nigeria. Study participants gave informed consent prior to their participation in the study. Every participant was made to understand that participation was entirely voluntary and that they could choose not to participate at any time. At the completion of the study and consistent with what was stated in the consent script, all intervention participants who listened to any part of the final call and control participants who listened to any part of the post-intervention survey received an incentive of a nominal amount of airtime credit equivalent to US$1.50 for their participation in the study.

## RESULTS

### Sociodemographic Characteristics of Study Participants

Table 1 compares the sociodemographic characteristics of the intervention and control groups. Overall, the average age was 26.8 years, and no significant difference existed between the intervention group (26.4 years) and the control group (27.0 years). The intervention and control groups were also equivalent in terms of marital status, education, and parity. In contrast, there were significant differences by religion. Specifically, a larger proportion of the intervention group was Muslims (65.6%) compared with the control group (65.6% vs. 57.2%, respectively; \( P < .05 \)).

### Exposure to the Intervention

User analytics data captured by the platform showed that the duration of listening differed between calls, which can be partially explained by variation in the content of the calls. The average listening times varied from less than 2 minutes to more than 10 minutes. Quiz calls 8 and 12 had the shortest average listening times. The longest average listening time was for the first call (10 minutes), which included the introduction and the pre-intervention survey, and call 17 (almost 11 minutes), which included the last quiz and the post-intervention survey. The average listening duration across all calls with program content was 5 minutes and 26 seconds.

The data further showed that the majority (96%) of the women in the intervention group listened to at least 1 complete episode of the drama—defined as a user listening to 100% of the episode. The episodes most likely to have been heard in their entirety by the study participants were the first 3 episodes, whereas exposure was relatively

---

The findings reported in this article were derived from both per-protocol and intention-to-treat DID.

The intervention and control groups were equivalent in terms of marital status, education, and parity but differed by religion.
lower for the last 3 of the 13 drama episodes. The average number of episodes completely heard was 7.2.

Table 2 describes variations in exposure to the various program components by key sociodemographic characteristics of the women in the intervention group. The data showed no differences in exposure to complete drama episodes by level of education or age group. On the other hand, exposure varied significantly by parity, religion, and marital status. On average, Muslim participants completed more drama episodes than Christian participants did (7.83 vs. 6.11, respectively; \( P < .001 \)). Similarly, ever-married women completed more episodes, on average, than their never-married peers (7.69 vs. 6.63, respectively; \( P < .05 \)).

Compared with the drama series episodes, exposure to the personal stories and the sample dialogues was lower. The participants in the intervention group listened to 2.59 complete personal stories and 1.15 sample dialogues, on average. Similar to what the data on the drama series episodes showed, exposure to the personal stories and the sample dialogues varied significantly by marital status, religion, parity, and education level (Table 2).

### Efficacy of the Program

The results of the DID analyses are provided in Table 3 and Table 4. The efficacy of the tool was initially assessed among the participants that answered the post-study questions (per-protocol analyses). A broader analysis was also conducted, focusing on the total sample and using intention-to-treat techniques. The results of the DID estimation adjusted for the participants’ age, education, religion, parity, and marital status (where appropriate). Notably, the women who participated in the post-study survey had a significantly higher level of exposure to the tool than their peers who did not participate in the post-study survey. For example, the women who participated in the post-study were exposed to a significantly higher mean number of drama series episodes (9.30) than their peers who did not participate in the post-study interview (5.75). In other words, the per-protocol effects reported below are probably indicative of what could be expected in the context of a high level of exposure to the tool by a wider audience.

Following is a presentation of the results for each outcome:

1. **Thoughts about desired family size.** The per-protocol analysis shows that at pre-study, women in the control group were more likely than their peers in the intervention group to have ever thought of their desired family size (42.5% compared to 33.0%). At the post-study, essentially no change had occurred in the control group. However, proportionally more women in the intervention group reported having ever given thought to the desired family size. The per-protocol DID estimate shows that the intervention led to a significant 43.2 percentage point increase in this indicator (Table 3). Results of the intention-to-treat analysis reveal a lower, albeit significant effect of 17.8 percentage points (Table 4).

2. **Confidence in one’s ability to discuss concerns about contraceptive methods with a provider.** According to the results of the per-protocol analysis, between pre-study and post-study, the proportion of participants confident in their ability to discuss concerns about contraceptive methods with a provider increased significantly in the intervention group (from 35.5% to 73.6%), whereas it declined conspicuously in the control group (from 59.5% to 36.1%). The reason for the huge decline among control group members is not clear. Results of the per-protocol DID
estimation reveal a 61.5 percentage point increase in this indicator attributable to the intervention (Table 3). In the intention-to-treat analysis, the DID was significant, but much smaller at 27.7 percentage points (Table 4).

3. **Discussion of desired family size with one’s spouse.** According to the results of the per-protocol analysis, the proportion of women that reportedly discussed desired family size with their spouse in the last 6 months was higher in the control group (74.6%) than in the intervention group (65.2%) at pre-study. At post-study, the indicator remained practically unchanged in the control group (66.7%) but increased in the intervention group (98.5%). The per-protocol DID estimate was 41.2 percentage points (Table 3), again indicating a significant positive effect of the intervention. Whereas the intention-to-treat estimate is much smaller (15.4 percentage points; Table 4), it remains nonetheless significant.

5. **Rejection of the misconception that contraceptive methods can harm the womb.** Results of the per-protocol analysis showed increased rejection of this misconception in the intervention group between pre-study (50.6%) and post-study (78.8%). In contrast, in the control group, proportionally fewer women (43.9%) at post-study rejected the misconception than at pre-study (64.1%). The per-protocol DID estimate stood large and significant at 48.4 percentage points (Table 3). The intention-to-treat estimate was smaller at 22.7 percentage points but still very significant (Table 4).

6. **Current use of modern contraceptive methods.** Whereas the use of modern contraceptive methods increased conspicuously

### Table 2. Mean Number of Smart Client Drama Episodes Completed, by Sociodemographic Characteristics, Intervention Group, Kaduna, Nigeria, 2017 (n=221)

<table>
<thead>
<tr>
<th>Sociodemographic Characteristics</th>
<th>Mean No. of Drama Episodes Completed</th>
<th>Mean No. of Personal Stories Completed</th>
<th>Mean No. of Sample Dialogues Completed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age group, years</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–24</td>
<td>6.89</td>
<td>2.33</td>
<td>1.81</td>
</tr>
<tr>
<td>25+</td>
<td>7.45</td>
<td>2.75</td>
<td>2.01</td>
</tr>
<tr>
<td><strong>Education level</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secondary or less</td>
<td>7.28</td>
<td>2.60</td>
<td>2.13*</td>
</tr>
<tr>
<td>Tertiary</td>
<td>7.17</td>
<td>2.58</td>
<td>1.53</td>
</tr>
<tr>
<td><strong>Marital status</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never married</td>
<td>6.63*</td>
<td>2.13*</td>
<td>1.25***</td>
</tr>
<tr>
<td>Ever married</td>
<td>7.69</td>
<td>2.93</td>
<td>2.44</td>
</tr>
<tr>
<td><strong>Religion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muslim</td>
<td>7.83***</td>
<td>2.95**</td>
<td>2.40***</td>
</tr>
<tr>
<td>Christian</td>
<td>6.11</td>
<td>1.89</td>
<td>1.04</td>
</tr>
<tr>
<td><strong>Parity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–1</td>
<td>6.65*</td>
<td>2.11*</td>
<td>1.25***</td>
</tr>
<tr>
<td>≥2</td>
<td>7.65</td>
<td>2.92</td>
<td>2.40</td>
</tr>
<tr>
<td>All</td>
<td>7.24</td>
<td>2.59</td>
<td>1.15</td>
</tr>
</tbody>
</table>

*P < .05; **P < .01; ***P < .001.
in the intervention groups (from 28.8% at pre-study to 63.6% at post-study), it remained at the same level in the control group (32.7%) at both time points. The estimated DID was 34.8 percentage points using the per-protocol approach (Table 3) and 14.8 percentage points using the intention-to-treat analysis (Table 4).

### Attrition

One major challenge encountered during the course of the study was the high attrition rate. A large number of women were recruited, but many of them did not engage at all with the platform and a significant number dropped out. Field workers recruited a total of 794 women (401 in the intervention group and 393 in the control group) into the study. This number included 641 originally recruited and 153 replacements. Of this number, only 559 (221 in intervention and 338 in control groups) took the Welcome Call and/or initiated the pre-intervention survey. The rate of noninitiation was higher among the women recruited into the intervention group (44.9%) than for their peers recruited into the control group (13.7%). The number of women that participated in the post-intervention survey was 92 for intervention and 158 for the control, a loss to follow-up (relative to study initiation) of 58.4% and 53.3%, respectively.

The differences in the pre-study sociodemographic characteristics between the women that participated in the post-study and their peers that were lost to follow-up are presented in Table 5. Among the women who participated in the intervention, the 2 groups were not significantly different in terms of age, marital status, and parity. The average age was 26.6 years in the group that participated in the post-intervention survey and 26.2 years in the lost-to-follow-up group. Mean parity was 2.43 for the post-study group compared with

### Table 3. Change in Selected Ideational and Behavioral Outcomes and Results of Differences-in-Differences (DID), per-Protocol Analyses, Kaduna, Nigeria, 2017

<table>
<thead>
<tr>
<th>Intervention Condition</th>
<th>Percent Reporting Outcome</th>
<th>DID Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-Study</td>
<td>Post-Study</td>
</tr>
<tr>
<td>Already thought of the number of children to have</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention group</td>
<td>33.0</td>
<td>77.5</td>
</tr>
<tr>
<td>Control group</td>
<td>42.5</td>
<td>43.8</td>
</tr>
<tr>
<td>Confident discussing family planning with provider</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention group</td>
<td>35.5</td>
<td>73.6</td>
</tr>
<tr>
<td>Control group</td>
<td>59.5</td>
<td>36.1</td>
</tr>
<tr>
<td>Discussed family size with spouse in last 6 months (currently married women only)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention group</td>
<td>65.2</td>
<td>98.5</td>
</tr>
<tr>
<td>Control group</td>
<td>74.6</td>
<td>66.1</td>
</tr>
<tr>
<td>Discussed contraceptive methods with spouse in last 6 months (currently married women only)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention group</td>
<td>46.4</td>
<td>75.8</td>
</tr>
<tr>
<td>Control group</td>
<td>43.0</td>
<td>49.7</td>
</tr>
<tr>
<td>Rejected the myth that contraceptive methods can hurt a woman’s womb</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention group</td>
<td>50.6</td>
<td>78.8</td>
</tr>
<tr>
<td>Control group</td>
<td>64.1</td>
<td>43.9</td>
</tr>
<tr>
<td>Using modern contraceptive method</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention group</td>
<td>28.8</td>
<td>63.6</td>
</tr>
<tr>
<td>Control group</td>
<td>32.7</td>
<td>32.7</td>
</tr>
</tbody>
</table>

*Total sample size (per protocol): intervention (n=92); control (n=158). All estimated models controlled for age, education, number of children ever born, and education.*
2.08 for the women lost to follow-up. In contrast, the 2 groups were significantly different in terms of religion and, to some extent, education. The differences by religion were such that 73.9% of the post-study group was Muslim compared with only about 60.1% of the lost to follow-up group. The difference by education was marginally significant: whereas 39.1% of the post-study group had tertiary education, only 28.1% of the women that were lost to follow-up did. In the control group, there were no significant sociodemographic differences between the women who participated in the post-study survey and their peers who were lost to follow-up.

A **IVR-based approach using drama is a viable option for promoting positive ideation related to family planning and increasing contraceptive use.**

**DISCUSSION**

Using longitudinal data, the analyses presented in this manuscript show that an IVR-based approach using drama is a viable option for promoting positive ideation related to family planning and increasing contraceptive use in Nigeria. The intention-to-treat DID results indicate that the intervention led to significant increases in contraceptive use and related ideational variables. This finding is consistent with prior reports on the use of mobile technologies in other settings. For example, in a study in Johannesburg, South Africa, Coleman and his colleagues assessed the effects of an SMS-based digital health intervention among pregnant women with HIV and found that the women that received the intervention had higher odds of obtaining the recommended amount of antenatal care visits, higher odds of normal vaginal delivery, and a lower odds of having a baby with low birth weight. Similarly, an SMS-based intervention in Nairobi, Kenya, was found to have led to increased exclusive breastfeeding and postpregnancy contraceptive use. In Zanzibar, Lund and her colleagues found that a mobile phone intervention increased skilled delivery attendance in urban but not in rural areas.

<table>
<thead>
<tr>
<th>TABLE 4. Change in Selected Ideational and Behavioral Outcomes and Results of Differences-in-Differences (DID), Intention-to-Treat Analyses, Kaduna, Nigeria, 2017³⁸ ³⁹</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intervention Condition</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Already thought of the number of children to have</td>
</tr>
<tr>
<td>Intervention group</td>
</tr>
<tr>
<td>Control group</td>
</tr>
<tr>
<td>Confident discussing family planning with provider</td>
</tr>
<tr>
<td>Intervention group</td>
</tr>
<tr>
<td>Control group</td>
</tr>
<tr>
<td>Discussed family size with spouse in last 6 months (currently married women only)</td>
</tr>
<tr>
<td>Intervention group</td>
</tr>
<tr>
<td>Control group</td>
</tr>
<tr>
<td>Discussed contraceptive methods with spouse in last 6 months (currently married women only)</td>
</tr>
<tr>
<td>Intervention group</td>
</tr>
<tr>
<td>Control group</td>
</tr>
<tr>
<td>Rejected the myth that contraceptive methods can hurt a woman’s womb</td>
</tr>
<tr>
<td>Intervention group</td>
</tr>
<tr>
<td>Control group</td>
</tr>
<tr>
<td>Using modern contraceptive method</td>
</tr>
<tr>
<td>Intervention group</td>
</tr>
<tr>
<td>Control group</td>
</tr>
</tbody>
</table>

³⁸Total sample size (intention to treat): intervention (n=220); control (n=339). All estimated models controlled for age, education, number of children ever born, and education.
Results from the evaluation of the Smart Client digital tool reveal some of the challenges of using digital health to disseminate health information. Some of these lessons have been highlighted in previous studies, while others are specific to the Smart Client digital tool. Such challenges might have contributed to the high attrition rate in the study. First, a high rate of noninitiation or nonengagement with the platform was present among the women who were recruited into the study. The noninitiation rate was considerably higher among the women recruited into the intervention group compared with those recruited into the control group. One possible explanation for the higher noninitiation rate for the intervention group is the intensity of the intervention. At recruitment, the intervention group was made aware there would be 17 calls, whereas the control group was informed that they would receive only 2 calls. A second challenge was that among the women in the intervention group who engaged with the platform, a large proportion did not complete the intervention.

Several other factors might have contributed to the high attrition rate, especially the noninitiation problem. For multiple reasons, a delay occurred between recruitment and when the calls began. Another issue, as reported by participants during follow-ups, was that the beginning of the calls sounded like automated calls from a marketer. For this reason, some participants did not complete the calls. Furthermore, some participants cited technical issues with the platform as being demotivating. These issues included calls not being sent out at the correct time due to network problems and SMS messages not being delivered to all participants who used a particular mobile network. The high attrition rate due to technical issues is not unique to the Smart Client tool, and studies have identified this problem in other digital health interventions. For example, a study in Ghana found that due to problems related to technical functionality, only about a quarter of the health messages sent by the Mobile Technology for Community Health’s mobile midwife program were received by the intended audience. Similarly, in Malawi, between 54% and 64% of SMS messages and between 27% and 38% of voice messages were successfully delivered to the intended audience. In contrast, registered users in the MomConnect mHealth program in South Africa received over 80% of the messages sent by the program. The developers of IVR platforms have also recognized the issue of attrition and noninitiation, especially in automated surveys, and for this reason, overrecruitment is commonly recommended.

### TABLE 5. Pre-Study Sociodemographic Characteristics of Study Participants, by Whether They Participated in the Post-Study Survey, Kaduna, Nigeria, 2017

<table>
<thead>
<tr>
<th>Sociodemographic Indicator</th>
<th>Participated in Post-Study Survey (n=92)</th>
<th>Lost to Follow-Up: Did Not Participate in Post-Study Survey (n=129)</th>
<th>Z (or t)/P for Difference Between Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intervention group</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age, years</td>
<td>26.6</td>
<td>26.2</td>
<td>0.653/.51</td>
</tr>
<tr>
<td>Currently married, %</td>
<td>55.4</td>
<td>52.3</td>
<td>0.453/.56</td>
</tr>
<tr>
<td>Tertiary education, %</td>
<td>39.1</td>
<td>28.1</td>
<td>1.716/.09</td>
</tr>
<tr>
<td>Muslim, %</td>
<td>73.9</td>
<td>60.1</td>
<td>2.123/.03</td>
</tr>
<tr>
<td>Mean parity</td>
<td>2.43</td>
<td>2.08</td>
<td>1.145/.25</td>
</tr>
<tr>
<td><strong>Control group</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age, years</td>
<td>27.2</td>
<td>26.8</td>
<td>0.777/.44</td>
</tr>
<tr>
<td>Currently married, %</td>
<td>57.3</td>
<td>57.0</td>
<td>0.063/.95</td>
</tr>
<tr>
<td>Tertiary education, %</td>
<td>35.4</td>
<td>27.9</td>
<td>1.470/.14</td>
</tr>
<tr>
<td>Muslim, %</td>
<td>59.7</td>
<td>54.1</td>
<td>1.052/.29</td>
</tr>
<tr>
<td>Mean parity</td>
<td>2.44</td>
<td>2.43</td>
<td>0.011/.99</td>
</tr>
</tbody>
</table>
With a study design that relied on the IVR platform to collect endline data, the high attrition rate could have significantly hampered the ability to make inferences. Fortunately, the study design used a prevalence of spousal communication (50%) that provided maximum variability to calculate the required sample size. The prevalence of spousal communication about family size among the intervention group turned out to be much lower than 50%, and it increased by more than the anticipated 15 percentage points at post-study. With these parameters, the reduced sample size at post-study still afforded us a 97.7% power to make inferences about the effects of the intervention. Furthermore, given the observed parameters, a repeated sample of 51 respondents is sufficient for a power of 80%. Nevertheless, the potential for noninitiation and high drop-out rate is a problem that should be accounted for when using mobile phone technology in the context of health behavior change interventions.

Some steps could be taken to minimize attrition and increase participation. For example, adequate measures should be put in place to minimize the delay between participants’ recruitment and the start of the intervention; ideally the delay should not be more than a few days. In addition, it is important that the calls do not start with a sound like an automated call from a marketer. It could be beneficial to inform participants at the time of recruitment what the opening segment of the calls will sound like. Participants should be asked or given assistance to program the platform phone number into their phone at the time of recruitment so they will recognize the incoming call. Furthermore, to avoid potential attrition due to technical issues with the platform, intensive testing should be conducted prior to wide-scale use. In the pilot test reported in this manuscript, participants were asked to select a convenient window of time for them to receive the calls. For each call, the system was automated to call back participants up to 6 times within this window of time. These accommodations added a layer of complication to the system that resulted in some groups not receiving calls or SMS messages. To avoid this problem, call scheduling options should be standardized for all participants.

Lessons Learned
This pilot test yielded some valuable lessons learned, including those related to tool development and implementation as well as those connected with the evaluation of effects.

- During the pretesting of the tool, participants expressed interest in more content. However, the analysis of listening patterns during this study indicates that most participants did not listen to additional segments of the calls. This gap between expressed interest and actual listening patterns has important implications for the design of the calls. Program implementers of future adaptations of the tool could consider shortening the content, eliminating segments, or splitting the segments into separate calls, so the calls are not so long. The modifications to the content, however, should be determined based on feedback from potential audience members that is ideally obtained through testing conducted by mobile phone delivery of the content to simulate the actual conditions of use.

- Although mobile phone penetration in developing countries has increased exponentially during the last few years, challenges remain with a mobile phone-based intervention due to the everyday issues faced by many owners and users of mobile phones. In the follow-ups with study participants, some commonly reported issues included sharing a phone with someone else, an inability to keep the phone regularly charged due to a lack of electricity, phones being lost or damaged, and switching phone numbers. These issues present various impediments due to the configuration of the platform and/or the design of the intervention. While these issues persist, some level of attrition must be assumed.

- Some study participants reported that they stopped listening due to their husband’s disapproval of their participation in the study. While this problem was not widespread, it does indicate the challenge of implementing a tool targeting women in locations where men make decisions for their wives; however, women in these locations are likely to be in greater need of the information included in this tool. Future interventions could potentially include husbands so they are able to listen to the content and discuss it with their wives.

- Unanticipated problems occurred with the IVR platform, especially with some features (e.g., flashing and SMS) not functioning correctly. When such issues took more time than expected to fix, they may have contributed to the attrition observed in this study. Testing of the platform was conducted prior to initiating
the study; however, more intensive testing of all the features of the platform and with all mobile networks should be conducted prior to widespread rollout of an intervention.

- No simple answer exists for the question of the cost effectiveness of the IVR approach compared to nondigital interventions. The costs associated with developing and implementing a tool such as Smart Client are highly dependent on various factors and decisions, but there are generally two main types of costs to consider: the costs of producing the content and the costs associated with the IVR platform.

  - Related to the content production, the HC3 team decided to work with a production house to record the content. This approach ensured that the quality of the audio content would be very high, both in terms of its presentation (with professional actors) and its sound quality (which is an important consideration when delivering via mobile phone). However, the content does not necessarily need to be professionally produced, and costs could be saved in this area. Another cost consideration related to the content development is the need to have content in multiple languages. This option is easy to accommodate on an IVR platform, which can be useful in locations where there are many local languages spoken; however, each language translation will add to the production costs.

  - The costs associated with setting up and maintaining an IVR platform can be very minimal (depending on the provider); however, the costs of the airtime minutes can add up very quickly especially when a call is long, which is another reason to reduce the amount of content. For the Smart Client testing and user study, the platform was set up to have all airtime costs charged to the project so that the calls would be free to users. This arrangement is the most common with use of IVR in developing countries not only because of the income level of audiences but also to encourage use of a tool that the audience may not be very familiar with. It is possible to set up an IVR platform that requires the listeners to pay a small amount to access the content. Willingness to pay a nominal fee was a concept explored by program implementers during pretesting but was not explored further in this study.

Lastly, costs of airtime minutes and SMS messages can be reduced through negotiations with mobile network operators; however, this can be a lengthy process and may necessitate guaranteeing that the tool will have a very large audience (i.e., customers with known identities). These factors point to an IVR tool similar to Smart Client being cost effective at a medium scale for a targeted audience and for a limited period of time; it should reach an audience that is large enough to justify the content production and airtime costs, but not so large that airtime costs over an unlimited amount of time are enormous.

### Evaluation of the Effects of the Tool

- Recruitment of participants into the study required specialized skills and a level of assiduity exceeding that for most other types of surveys. The recruiters needed to understand that the study participants would have to commit to receiving multiple program calls and stay in the program for up to 3 months. Moreover, recruitment required testing potential participants’ numeracy skills and Hausa linguistic skills. Failure on the part of recruiters to completely apply recruitment guidelines might have contributed to the initial failure of some participants to engage with the platform and the high level of dropout along the way.

- Merging data from the various program calls with data collected during recruitment was difficult at best. This problem was due to the fact that the participant’s telephone number, which was intended to be used as the unique identifier, was not consistent across calls. Moreover, the format used for recording the telephone number was not consistent between the data collected at recruitment and the data collected through the program calls. Harmonizing the formatting required considerable data manipulation skills and was time consuming. As a result, matching some cases across program calls was not possible. While it is impossible to avoid the situation of people using different cell phones across calls, evaluation and program teams as well as platform technicians should work together to ensure that the same formatting style is used for essential data fields across multiple data sources. For example, including or excluding the country code in a telephone number makes a lot of difference for the ability to match data from various sources.
• In this study, due to the high attrition rate, the post-intervention survey came close to not having a sufficient sample to permit inferences being made. Evaluation of future adaptations of the tool using randomized control trial or cluster-randomized control trial should anticipate and adequately plan for a higher than usual attrition rate.

**CONCLUSION**

The Beta Life version of the Smart Client digital health tool offers a potentially effective approach for promoting positive contraceptive attitudes and encouraging women to adopt a modern contraceptive method. The tool could contribute to increasing contraceptive prevalence and reducing unmet need for contraceptives in Nigeria. Future adaptations of the tool should address the limitations connected with the number and length of program calls, its requirement for considerable numeracy skills in low literate settings, and the issues related to recruitment and initiation. Additional information about the tool, including full scripts and an Adaptation Guide can be found at [https://healthcommcapacity.org/technical-areas/family-planning/smart-client-smart-couples/](https://healthcommcapacity.org/technical-areas/family-planning/smart-client-smart-couples/).

**Acknowledgments:** The authors acknowledge the management and data collectors from Market Audits and Research Services, under the leadership of Desmond Nweke, for helping to collect the data analyzed in this manuscript. Our appreciation also goes to the women in Kaduna city who participated in the study.

**Funding:** This research was made possible through the generous support and funding from the United States Agency for International Development (USAID) through the Health Communication Capacity Collaborative Project under Cooperative Agreement #AID-OAA-A-12-00058.

**Disclaimer:** The opinions expressed in this manuscript are the authors’ and do not reflect the view of USAID or of Johns Hopkins University.

**Competing Interests:** None declared.

**REFERENCES**


20. Ejembi CL, Dahiru T, Aiyiku AA. Contextual Factors Influencing Modern Contraceptive Use in Nigeria. Rockville, MD: ICF


Adding a Question About Method Switching to the Method Information Index Is a Better Predictor of Contraceptive Continuation

Aparna Jain, a Kumudha Aruldas, b Elizabeth Tobey,a Arupendra Mozumdar, c Rajib Acharyac

Adding the question “Were you told about the possibility of switching to another method if the method you selected was not suitable?” to the Method Information Index (MII) was associated with better contraceptive continuation. This MIIplus variable includes another domain of quality of care, and thus better reflects voluntary contraceptive use and continuation.

ABSTRACT

Introduction: The Method Information Index (MII) is 1 of 18 core indicators used to monitor progress toward achieving Family Planning 2020’s goal of 120 million more women using a modern method of family planning by 2020. The 3 questions of the MII are intended to measure informed choice at method initiation. Although routinely used in the Demographic and Health Surveys and the Performance Monitoring and Accountability 2020 project in cross-sectional household surveys, the MII may not adequately reflect all key aspects of quality of care or predict contraceptive continuation. In the current study, a question was added to the MII regarding the possibility of switching to a different contraceptive method if the current method is not suitable. The revised MII is referred to as MIIplus.

Methods: A total of 2,699 married women aged 15–49 who started a new episode of use of intrauterine device, injectable, or oral contraceptive pills between December 2016 and October 2017 were followed for 1 year in India and interviewed at method start and 3, 6, and 12 months later. Of these women, 2,267 were interviewed 3 months later and included in the analysis. Using 3 Cox proportional hazard models, we estimated hazard ratios for risk of discontinuation, based on the MII, MIIplus, and a recategorization of MIIplus into a 3-category variable.

Results: The modern method continuation rate 100 days (~3 months) later was 91% overall. Women who received the information in MIIplus were more likely to continue using a method at 100 days (95%) compared to those who received information covered in the MII (82%) or less than 3 components of the MII (89%) (P < .001). Women who received all components in the MIIplus were 69% (adjusted hazard ratio, 0.31; 95% confidence interval: 0.17 to 0.61) less likely than those who received information in the MII to discontinue using a modern method 100 days later. Discontinuation was not significantly different between women who received information on less than the 3 components of the MII compared to the complete MII.

Conclusion: We recommend including the question about the possibility of switching to another family planning method in routine measurement because it better predicts contraceptive continuation than the MII alone and ensures that another domain of quality of care is reflected in the measurement. When programs provide information on the possibility of switching, women are better informed about voluntary family planning choice and their options to continue family planning when a method is not suitable.

INTRODUCTION

The Method Information Index (MII) was created as an indicator of informed choice by Family Planning 2020 (FP2020). It is 1 of 18 core indicators used to monitor progress toward achieving the goal of 120 million additional contraceptive users among women with an unmet need for family planning by 2020.1 Routinely collected in the Demographic and Health Surveys (DHS) for many years, the MII comprises 3 questions related to women’s reports about the information received at the time of method adoption. It is intended to assess the presence of informed choice and includes items from 2 domains of quality of care: (1) information given to clients and solicitation of information from clients for appropriate method selection (information exchange about method selection), and (2) information given to clients by the provider about the method selected (effective use of method selected).2
Method Switching Question and the Method Information Index

The MII uses the following 3 questions:

1. Were you informed about other methods of family planning?
2. Were you informed about possible side effects or problems you might have with the method?
3. Were you told what to do if you experience any side effects or problems? (asked among those who were told about side effects)

Women who respond yes to all 3 questions are considered to have received full information, based on the MII.

In cross-sectional household surveys, the MII questions are included in the DHS and in the Performance Monitoring and Accountability 2020 (PMA2020) project. In the DHS, the MII questions are asked of current family planning users who adopted the method within 5 years of the survey. In PMA2020 surveys, the MII questions are asked of current and recent users who adopted a method in the past 12 months. MII data have been analyzed to compare differences among regions and countries and to study variations by female respondents’ characteristics and country-level changes over time.

Although these questions have been asked in the DHS for many years, the combination of the 3 questions as the MII is relatively new and thus, much is still unknown about the MII. One question is whether it adequately reflects all 4 domains of quality of care: (1) respectful care; (2) information exchange about method selection; (3) effective use of method chosen; and (4) continuity of contraceptive use and care. Another question is whether receiving all pieces of information covered by the MII relates to contraceptive continuation.

Using longitudinal data collected from 2,699 women initiating use of reversible modern contraceptive methods in India, we investigated whether adding the following question to the MII, “Were you told about the possibility of switching to another method if the method you selected was not suitable?”, was associated with a reduced risk of modern method discontinuation 100 days later (~3 months) compared with receiving information in the MII alone.

METHODS

Study Areas

The Evidence Project, led by the Population Council, conducted a longitudinal study of married women aged 15–49 in 2 states in India, Haryana and Odisha, who began a new episode of use of a reversible contraceptive method. These 2 states were selected in consultation with the Ministry of Health.

The total fertility rate in India is 2.2 children per woman, and slightly more than 1 in 2 (54%) married women use any type of contraceptive method to space or limit childbearing. Use of any contraceptive method is greater in Haryana (64%) and Odisha (57%) than in the all-India proportion. The total fertility rate in Odisha declined from 2.9 children per woman in 1992–1993 to the replacement level of 2.1 children in 2015–2016, while a more rapid decline was seen in Haryana (4.0 children per women in 1992–1993 to 2.1 in 2015–2016). In Odisha, 45% of married women between 15 and 49 years old used a modern contraceptive method in 2015–2016, with the pill being the most popular reversible method (12%). In comparison, Haryana had slightly more users of modern contraceptive (59%), with condoms being the most popular reversible contraceptive method (12%). Discontinuation of episodes of use within 12 months was 48% for modern reversible methods in Odisha compared with 41% in Haryana.

This study reflects the current situation related to family planning counseling in the selected areas. No intervention or provider trainings were conducted to improve counseling.

Data

A new episode of use is defined as a new user of family planning or a past user of family planning who was not using a method right before the method selected at enrollment. These users of the intrauterine device (IUD) (interval and postpartum), injectable, or oral contraceptive pills (OCPs) began a new episode of use between December 2016 and October 2017. They were interviewed within 1 month of the start of this new episode of use (enrollment interview) and at 3, 6, and 12 months following the first interview. The overall purpose of this study was to explore the contraceptive use dynamics of reversible contraceptive users including barriers and facilitators to contraceptive continuation, discontinuation, and method switching. One facilitator/barrier of contraceptive use dynamics explored was quality of care received at the time of method adoption.

Respondents were enrolled from government and private health facilities as well as through accredited social health activists (ASHAs), who are frontline health workers at the community level. In Odisha, at the suggestion of district chief
medical officers, postpartum IUD (PPIUD) users were mostly enrolled at selected government health facilities from the labor and delivery wards. In some cases, ASHAs who were identified by the facility managers, enrolled PPIUD users. ASHAs also enrolled interval IUD and OCP users. Injectable users were enrolled at NGO facilities. In Haryana, all respondents were enrolled by ASHAs at the community level.

Face-to-face interviews were conducted at the facility or at the respondent’s household, depending on her preference. Study investigators were trained in the study objectives, questionnaires, and informed consent procedures at the beginning of the study and at the start of each follow-up interview. A total of 2,699 women were enrolled in the study and 2,306 were successfully re-interviewed 3 months later, for a loss to follow-up of 14.6%. An additional 23 respondents were excluded from the analysis because they stopped using contraception to become pregnant, and 16 were excluded because of inconsistent dates.

The current study uses data from enrollment and 3-month follow-up interviews for 2 reasons. First, we hypothesized that quality of care received at the time of method adoption would directly affect contraceptive use dynamics in a shorter time frame than a longer one. Second, we hypothesized that women who experience side effects would be more likely to experience them within the first 3 months of use; thus the risk of discontinuation might be greatest during this time period.

Written consent was obtained from all respondents at the enrollment interview and each follow-up interview. The study received ethical approval from the Population Council Institutional Review Board, the Government of Odisha, and district authorities in selected districts in Haryana.

Dependent Variable
In this study, the dependent variable is the time in days until the respondent discontinued modern contraceptive use. The observation period is from enrollment until the 3-month follow-up interview or approximately 100 days. All nonusers and traditional method users at the 3-month follow-up interview were considered modern contraceptive discontinuers. Modern contraceptive continuers included users of the method adopted at enrollment as well as method switchers who reported using a modern method at the 3-month follow-up interview that was not the method initiated at enrollment.

Key Independent Variables
Method information indicators are the key independent variables. They include the MII and method information index plus information received about the possibility of switching methods (MIIplus). In addition, MIIplus was recategorized into a 3-category variable and tested as a separate independent variable.

MII
The MII is composed of the 3 questions presented earlier (in the introduction). The MII combines the responses into a dichotomized variable, with women who reported yes to all 3 questions being coded as 1, and women who reported yes to fewer than 3 of these questions being coded as 0.

MIIplus
The MIIplus variable adds the question “Were you told about the possibility of switching to another method if the method you selected was not suitable?” to the 3 questions of the MII. This question was asked of all women and not just those who reported receiving information about possible side effects. First, this variable was dichotomized, with women who reported yes to all 4 questions being coded as 1 and women who reported yes on fewer than 4 of these questions being coded as 0. A 3-category variable disaggregating the 0 category of the MIIplus variable was then constructed to create the MIIplus variation. In this variation, the variable was coded as follows: women who reported yes on fewer than 3 questions of the MII were coded as 0; women who reported yes on all 3 questions of the MII (but did not receive information about the possibility of switching methods) were coded as 1; and women who reported yes on all questions of MIIplus were coded as 2.

Additional Independent Variables
Several covariates were included in the multivariate models based on theoretical and empirical importance of contraceptive use including age, education, religion, number of living children, method selected at enrollment, previous modern method use, and state.

Data Analysis
Descriptive statistics were calculated for respondent characteristics and dependent and key independent variables. Survival analysis was applied from the date of enrollment to the date a respondent reported discontinuing a modern contraceptive
method. If a respondent was lost to follow-up at the 3-month interview or discontinued using any method because of a desire to become pregnant, the respondent was censored. Kaplan-Meier survival functions were used to estimate continuation rates among the various method information measures. These functions were also used to estimate modern method continuation rates by MIIplus for each method adopted at enrollment. That is, estimates of modern method continuation rates included respondents who were using the same method adopted at enrollment 100 days later and those who switched to another modern method in this period. This shifts the unit of analysis from a specific method to an individual.

Cox proportional hazard models were used to estimate hazard ratios for risk of modern contraceptive discontinuation. Proportional hazard assumptions were checked based on the scaled Schoenfeld residuals. Three hazard models were run, one for each key independent variable. The Akaike information criterion (AIC) was applied to assess model fit between Models I (MII) and II (MIIplus) only because Model III (MIIplus variation) is simply a test of an association, while Models I and II reflect the ways in which the MII and MIIplus are derived. All statistical analyses were conducted in Stata Version 13 (StataCorp, 2013).

## RESULTS

### Respondent Characteristics

Table 1 presents respondent characteristics at enrollment. Most married women enrolled in the study were under the age of 30 (74%), had attended at least primary school (78%), and were Hindu (84%). Nearly all women (99%) had at least 1 child, and a quarter had 3 or more children. Thirty-eight percent had used a modern method previously. In our sample, 40% were OCP users, 39% were IUD users (15% PPIUD and 24% interval IUD), and 21% injectable users.

### Method Information Measures

The distributions of the various method information measures are also presented in Table 1. Thirty-seven percent of respondents received complete information in the MII while 34% received complete information in the MIIplus. When MIIplus was disaggregated into 3 categories (MIIplus variation), only 3% received complete information in the MII.

Overall probability of continuing modern method use 100 days later within this sample was 91% (data not shown). Table 2 shows the probability of modern contraceptive continuation at 100 days by the MII, MIIplus, and MIIplus variation. Among respondents who reported receiving complete information in the MII, 94% continued to use a modern contraceptive method 100 days later compared with 89% among those who responded “yes” to fewer than 3 questions of the MII (P<.001) (Figure 1). Ninety-five percent of respondents who reported receiving complete information in MIIplus continued to use a modern contraceptive method 100 days later compared with 89% who reported receiving less information (P<.001) (Figure 2). Figure 3 shows the Kaplan-Meier survival curves at 100 days for modern contraceptive continuation by the 3 categories of the MIIplus variation. Women who received all information in MIIplus (95%) were more likely to continue using a modern contraceptive at 100 days compared with those who received information covered by the MII (82%) or less than 3 components of the MII (89%) (P<.001).

### Method Information Indicators and Risk of Modern Method Discontinuation

Crude and adjusted hazard ratios (HRs) of the risk of contraceptive discontinuation are presented in Table 3. Model I shows that the adjusted HR for discontinuation among women who received complete information based on the MII was 0.65 (95% confidence interval [CI]=0.47 to 0.91) compared with women who received information on less than 3 components, adjusted for age, education, religion, number of living children, previous modern method use, method selected at enrollment, and state. Women who received information on all 4 components of MIIplus were less likely to discontinue 100 days later than those who received information on less than the 4 components (Model II adjusted HR=0.53; 95% CI=0.37 to 0.76). The AIC for MII (Model I) was higher than the AIC for MIIplus (Model II), suggesting a better fit of the data for MIIplus.

Model III shows the adjusted HRs for discontinuation using the MIIplus variation variable. Women who received all information included in MIIplus were 69% less likely to discontinue using a modern method 100 days later compared with women who received the information in the MII (adjusted HR=0.31; 95% CI=0.17 to 0.61). No difference was observed in discontinuation among


women who received information on less than 3 components of the MII compared to the complete MII.

**Modern Method Continuation by Enrollment Method and MIIplus**

Figure 4, Figure 5, and Figure 6 present the Kaplan-Meier survival curves for modern method continuation by MIIplus among IUD, injectable, and pill users at enrollment. The probability of continuing a modern contraceptive method at 100 days for IUD (both PPIUD and interval IUD), injectable, and pill users who received complete information in MIIplus was greater than among women who did not receive complete information in MIIplus. Continuation among IUD users at enrollment who received all information in MIIplus was greater than among women who did not receive complete information in MIIplus. Continuation among IUD users at enrollment who received all information in MIIplus was 96% compared with 86% among IUD users who received information covered by MII but not about method switching (Figure 4). Similar relationships were seen for the injectable (Figure 5) and the pill (Figure 6), with injectable and pill users who received information in the MII and about the possibility of method switching being more likely to continue using 100 days later (89% for the injectable and 96% for the pill). In addition, for injectable users, most discontinuation occurred around the 100-day mark for women who received full information in the MIIplus and women who received information on less than all components included in MIIplus. Since the injectable is a 3-month method, active discontinuation would most likely occur around the time of reinjection.

**DISCUSSION**

This study showed that women who report receiving full information included in MIIplus had

---

**TABLE 1. Respondent Characteristics and Method Information Measures at Enrollment (N=2,699)**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td></td>
</tr>
<tr>
<td>≤24</td>
<td>1,066 (39.5)</td>
</tr>
<tr>
<td>25–29</td>
<td>935 (34.6)</td>
</tr>
<tr>
<td>≥30</td>
<td>698 (25.9)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>605 (22.4)</td>
</tr>
<tr>
<td>Primary</td>
<td>335 (12.4)</td>
</tr>
<tr>
<td>Middle</td>
<td>386 (14.3)</td>
</tr>
<tr>
<td>Secondary</td>
<td>798 (29.6)</td>
</tr>
<tr>
<td>Higher secondary</td>
<td>575 (21.3)</td>
</tr>
<tr>
<td>Religion</td>
<td></td>
</tr>
<tr>
<td>Hindu</td>
<td>2,272 (84.2)</td>
</tr>
<tr>
<td>Muslim</td>
<td>418 (15.5)</td>
</tr>
<tr>
<td>Other</td>
<td>9 (0.3)</td>
</tr>
<tr>
<td>No. of living children</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>23 (0.9)</td>
</tr>
<tr>
<td>1</td>
<td>1,118 (41.4)</td>
</tr>
<tr>
<td>2</td>
<td>891 (33.0)</td>
</tr>
<tr>
<td>3 or more</td>
<td>667 (24.7)</td>
</tr>
<tr>
<td>Previous modern method use</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1,037 (38.4)</td>
</tr>
<tr>
<td>No</td>
<td>1,662 (61.6)</td>
</tr>
<tr>
<td>Method selected at enrollment</td>
<td></td>
</tr>
<tr>
<td>OCPs</td>
<td>1,066 (39.5)</td>
</tr>
<tr>
<td>PPIUD</td>
<td>412 (15.3)</td>
</tr>
<tr>
<td>Interval IUD</td>
<td>640 (23.7)</td>
</tr>
<tr>
<td>Injectable</td>
<td>581 (21.5)</td>
</tr>
<tr>
<td>State</td>
<td></td>
</tr>
<tr>
<td>Haryana</td>
<td>908 (33.6)</td>
</tr>
<tr>
<td>Odisha</td>
<td>1,791 (66.4)</td>
</tr>
<tr>
<td>MII &lt;3 questions</td>
<td>1,711 (63.4)</td>
</tr>
<tr>
<td>MII All 3 questions</td>
<td>988 (36.6)</td>
</tr>
<tr>
<td>MIIplus &lt;4 questions</td>
<td>1,779 (65.9)</td>
</tr>
<tr>
<td>MIIplus All 4 questions</td>
<td>920 (34.1)</td>
</tr>
</tbody>
</table>

**TABLE 1. Continued**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIIplus variation</td>
<td></td>
</tr>
<tr>
<td>&lt;3 questions of MII</td>
<td>1,711 (63.4)</td>
</tr>
<tr>
<td>MII</td>
<td>68 (2.5)</td>
</tr>
<tr>
<td>MIIplus</td>
<td>920 (34.1)</td>
</tr>
</tbody>
</table>

Abbreviations: IUD, intrauterine device; MII, Method Information Index; MIIplus, Method Information Index including the question, “Were you told about the possibility of switching to another method if the method you selected was not suitable?”; OCPs, oral contraceptive pills; PPIUD, postpartum intrauterine device.
TABLE 2. Kaplan-Meier Estimates of 100-Day Continuation of Modern Contraception (n=2,267)

<table>
<thead>
<tr>
<th>Method Information Indicators</th>
<th>% (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MII</td>
<td></td>
</tr>
<tr>
<td>&lt;3 questions</td>
<td>89.3 (87.5, 90.8)</td>
</tr>
<tr>
<td>All 3 questions</td>
<td>93.7 (91.7, 95.2)</td>
</tr>
<tr>
<td>MIIplus</td>
<td></td>
</tr>
<tr>
<td>&lt;4 questions</td>
<td>89.0 (87.3, 90.5)</td>
</tr>
<tr>
<td>All 4 questions</td>
<td>94.7 (92.7, 96.1)</td>
</tr>
<tr>
<td>MIIplus variation</td>
<td></td>
</tr>
<tr>
<td>&lt;3 questions of MII</td>
<td>89.3 (87.5, 90.8)</td>
</tr>
<tr>
<td>MII</td>
<td>82.3 (70.3, 89.8)</td>
</tr>
<tr>
<td>MIIplus</td>
<td>94.7 (92.7, 96.1)</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; MII, Method Information Index; MIIplus, Method Information Index including the question, “Were you told about the possibility of switching to another method if the method you selected was not suitable?”

The MIIplus variable behaves similarly irrespective of method.

TABLE 2. Kaplan-Meier Estimates of 100-Day Continuation of Modern Contraception (n=2,267)

The Kaplan-Meier survival estimates of MIIplus by method adopted at enrollment suggest that the MIIplus variable behaves similarly irrespective of method. The probability of modern contraceptive continuation among women who adopted the IUD, injectable, or pill and reported receiving complete information in MIIplus at method adoption was greater compared with those who had received information on less than 4 components included in MIIplus.

The possibility of switching methods may have been captured in the question, “Were you told what to do if you experience side effects?” In this study, a follow-up question (“What were you told?”) was asked of respondents who reported being counseled on what to do if they experienced side effects. No respondent reported that they were told about the possibility of switching. Furthermore, typically the question about what to do about side effects was only asked of respondents who reported that they were told about side effects. Information about the possibility of switching should be offered to all women and not necessarily linked to the experience of side effects because women discontinue contraceptive methods for numerous reasons, including cost and access. Also, the impact of physical side effects on women’s lives and relationships could also affect decision making around method continuation, which may not necessarily be discussed in family planning counseling.

FIGURE 1. Kaplan-Meier Survival Curves of Contraceptive Continuation by MII

Abbreviation: MII, Method Information Index.
The MII has been explored in 2 other recent studies. A longitudinal study among family planning clients attending private franchise clinics in Pakistan and Uganda showed that women who did not receive full information according to the MII had a significantly increased hazard of contraceptive discontinuation 6 months after method adoption in Pakistan. The association was also observed among family planning clients in Uganda, but it was not significant. While the results presented here confirm these findings in public health facilities and through community-based distribution, the current study suggests that the observed relationship between the MII and discontinuation 3 months later may have been due to the fact that a large proportion of women also reported receiving information about method switching.
Based on the same datasets from Pakistan and Uganda, follow-up questions were asked of family planning clients after the MII questions to assess consistency in reporting. Follow-up questions included asking respondents to name another family planning method, specific side effects associated with the method received, and specific actions to take if a side effect was experienced. While the authors concluded that they observed significant decreases in the MII scores when they adjusted for consistency of the follow-up questions, consistency of 2 individual MII questions seemed high: (1) informed about other methods and naming another method, and (2) informed what to do if a side effect was experienced.

### TABLE 3. Crude and Adjusted Hazard Ratios of Contraceptive Discontinuation 100 Days Later (n=2,267)

<table>
<thead>
<tr>
<th>Models of Method Information Indicators</th>
<th>HR (95% CI)</th>
<th>AHRa (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model I: MII</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;3 questions</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>All 3 questions</td>
<td>0.59*** (0.42, 0.80)</td>
<td>0.65* (0.47, 0.91)</td>
</tr>
<tr>
<td>Model II: MIIplus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;4 questions</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>All 4 questions</td>
<td>0.48*** (0.33, 0.68)</td>
<td>0.53*** (0.37, 0.76)</td>
</tr>
<tr>
<td>Model III: MIIplus variation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;3 questions of MII</td>
<td>0.57 (0.31, 1.06)</td>
<td>0.56 (0.30, 1.04)</td>
</tr>
<tr>
<td>MII</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>MIIplus</td>
<td>0.28*** (0.14, 0.55)</td>
<td>0.31*** (0.17, 0.61)</td>
</tr>
</tbody>
</table>

Abbreviations: AHR, adjusted hazard ratio; CI, confidence interval; HR, hazard ratio; MII, Method Information Index; MIIplus, Method Information Index including the question, “Were you told about the possibility of switching to another method if the method you selected was not suitable?”

* P ≤ 0.05; ** P ≤ 0.01; *** P ≤ 0.001.

a Adjusted for age, education, religion, number of living children, previous modern method use, method selected at enrollment, and state.

### FIGURE 4. Kaplan-Meier Survival Curves of Contraceptive Continuation by MII With the Added Question About Method Switching (MIIplus) Among IUD Users at Enrollment

Abbreviations: IUD, intrauterine device; MII, Method Information Index.
are experienced and knowing the specific actions to take. Consistency was less likely when women were asked to name a specific side effect associated with the method they received. Additional work in this area is needed to validate a measure that accurately captures side effect counseling that women receive and the information that they take from it.

Another recent analysis compared measures of counseling on side effects captured in 3 data sources: (1) family planning clients’ report of information received reported in the household survey of the DHS; (2) direct observations of client-provider interactions captured in the Service Provision Assessment (SPA); and (3) client exit interviews collected in the SPA. When comparing the SPA data sources, the author found that clients tended to overestimate information received in comparison to direct observations.
While biases like courtesy bias and acquiescence bias may contribute to this overestimate, information that clients take away from their family planning consultations should be the focus of quality of care measurement because this information will likely influence their subsequent behaviors.

The results of the current study suggest that discontinuation of modern methods 100 days after initiation is lower for MIIR compared with the MII alone in this sample population. MIIR is a marker of good contraceptive counseling and includes information on an additional element of quality of care—continuity of use and follow-up. This information may help women continue to use modern contraception even if they discontinue their initially adopted method. For example, if a woman is told what to do if she experiences side effects but the recommended course of action is ineffective, then she may discontinue using contraception altogether because she did not know that she had the option to switch to a different method.

While MIIR is a better measure of quality of care compared with the MII (with the addition of another domain of quality of care) for voluntary family planning choice, it still falls short of reflecting all 4 domains of quality of care. MIIRplus, however, could be used in cross-sectional and population-level surveys to obtain trends over time and across countries because it has been shown to predict lower levels of discontinuation compared with the MII. Additional measures that reflect all domains of quality of care have been proposed for routine monitoring of programs.11

As suggested elsewhere,3 adding the switching question for reversible contraceptive method users complements the question asked of sterilization users—whether they were told it was permanent—so that MIIRplus could be based on 4 questions collected for all contraceptive users. The use of MIIRplus as an indicator may encourage family planning programs and providers to move away from emphasizing method-specific continuation to continuation of any modern family planning (among women with a need for family planning). Furthermore, with the use of MIIRplus in programs, women who find that their current method is no longer suitable at any time or for any reason may be more likely to switch to another contraceptive method.

Limitations

The current study has several limitations. Most Indian women in this study who received full information according to their responses to questions in the MII also received information about the possibility of switching, suggesting that in India, providers have a tendency to routinely offer this information in family planning counseling. Future research in contexts in which this information may not be routinely given will help to further elucidate the relationship between being told about switching and modern contraceptive continuation.

A total of 432 respondents were excluded from these analyses because they were lost to follow-up, discontinued method use because they wanted to get pregnant, or had data inconsistencies. Logistic regression comparing these individuals with those who were included in these analyses (n=2,267) by respondent characteristics revealed no differences between these 2 groups for most characteristics (data not shown). However, 2 characteristics—religion and number of living children—were significantly different for respondents included and excluded from the analyses. Muslims compared with Hindus were significantly less likely to be excluded from these analyses, as were women with 2 or more children compared with those with 1 child. Because the final hazard models were adjusted for respondent characteristics, these differences were unlikely to have had a substantial effect on the conclusions.

Additional study limitations were that participants were not enrolled at the same time and the enrollment strategy varied by state and family planning method. IUD users were enrolled at public district hospitals, while eligible OCP users were identified through ASHAs at the community level. At the time of the study, the Indian government had committed to expanding method choice with the introduction of the injectable through public facilities, but rollout had not begun. Consequently, injectable users were recruited from NGO facilities. Thus, any differences observed in the quality of care received by clients for specific methods is likely due to the facility where the services were obtained rather than the method itself. Respondents in Haryana were enrolled in the study several months after enrollment began in Odisha. Future research should consider including the question about switching and replicating the analysis presented here to build the evidence base for the MIIRplus.

Conclusion

This study demonstrates that modern method discontinuation 3 months after method initiation is lower when women receive information in the
MIIplus compared to the MII alone. By adding the question about method switching to the MII, MIIplus better reflects quality of care. Furthermore, shifting the outcome from method-specific discontinuation to modern method discontinuation allows the unit of analysis to reflect individual-level behaviors. Policymakers, programmers and donors who rely on the MII for national and sub-national estimates may be better served by using MIIplus to monitor quality of care in family planning services.

Acknowledgments: The authors thank Anrudh K. Jain for his useful comments on a previous version of the manuscript. The authors thank research assistants who collected the data in India and to the women who participated in the study and shared their time and experiences.

Funding: The Evidence Project is made possible by the generous support of the American people through the United States Agency for International Development (USAID) under the terms of cooperative agreement No. AID-OAA-A-13-00087.

Disclaimer: The contents of this manuscript are the sole responsibility of the authors and do not necessarily reflect the views of USAID or the United States Government.

Competing Interests: None declared.

REFERENCES

Peer Reviewed

Received: January 11, 2019; Accepted: April 16, 2019


© Jain 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-19-00028
Associations Between Practices and Behaviors at the Health Facility Level and Supply Chain Management for Antiretrovirals: Evidence from Cameroon, Namibia, and Swaziland

Diana Bowser,a Laura Krech,b David Mabirizi,c Angela Y. Chang,d David Kapaon,e Thomas Bossertf

Using antiretrovirals (ARVs) as tracer products, we identified the following key practices that may affect supply chain management at the facility level: order verification, actions taken when stock is received, changes in prescription and dispensing due to ARV stock-out, actions to ensure patient adherence, and communication with other affiliated facilities and higher-level supply chain management. We propose a set of indicators to measure these practices.

ABSTRACT

Background: While measuring, monitoring, and improving supply chain management (SCM) for antiretrovirals (ARVs) is understood at many levels of health systems, a gap remains in the identification and measurement of facility-level practices and behaviors that affect SCM. This study identifies practices and behaviors that are associated with SCM of ARVs at the hospital level and proposes new indicators for measurement.

Methods: We performed an in-depth literature review to identify facility-level practices and behaviors and existing indicators that are associated with SCM. We used the United States Agency for International Development’s 2013 National Supply Chain Assessment Toolkit to define 7 supply chain function areas to frame the study. Qualitative, semistructured key informant and focus group interviews were conducted in hospitals with health professionals from Cameroon, Namibia, and Swaziland to understand facility-level practices and behaviors.

Results: Using the results from 54 key informant and focus group interviews from 12 hospitals, we identified 30 practices and behaviors that may affect ARV SCM at the facility level. The following practice areas were particularly associated with SCM: order verification, actions taken when ARV stock is received, changes in prescription and dispensing due to ARV stock-out, actions to ensure patient adherence, and communication with other affiliated facilities and higher-level SCM. We subsequently developed measurable indicators for future research.

Conclusion: This study characterizes facility-level practices and behaviors that can affect ARV SCM. It also identifies gaps in their measurement. While this study uses ARVs as a tracer medicine to understand gaps in practices at the facility level, many of the findings are more broadly applicable to other medicines in an integrated setting. This study provides real-world evidence and the groundwork for further research to characterize the link between 30 facility-level practices and behaviors and ARV SCM at the facility and central levels.
organizations. Despite these efforts, supply chain management (SCM) remains a weak part of the national health system in many countries, where the delivery of ARVs is not yet sustainable or fully transitioned to the MOH, leaving the risk of stockout for ARVs high.

Effective and efficient communication between central, regional, and local health facility levels as well as quantification, secure transport, and quality assurance of medical products are all required parts of a functional supply chain. Although much is understood about how to measure, monitor, and improve ARV SCM at the central and regional levels of a health system, less has been done to identify and measure facility-level practices and behaviors that can affect SCM. In this article, we define facility-level SCM as activities related to managing and ordering inventory and monitoring the performance of facilities at the point of service delivery carried out by health care providers, managers, and administrators involved in prescribing and dispensing. These facilities may include public and private hospitals and health facilities, clinics, pharmacies, drug stores, and other outlets. Practices are defined as actions taken (or lack of actions taken) within facilities by health care workers and/or managers that affect SCM. Behaviors are defined as the leadership or management styles as well as the quality of relationships and communication between health care workers and managers at the facility level and those working in higher-level SCM management positions.

Comprehensive, well-designed reports and tools already exist with measurable indicators across various aspects of the supply chain. They include Management Sciences for Health’s (MSH) Managing Drug Supply-3 (MDS-3): Managing Access to Medicines and Health Technologies; Logistics Indicator Assessment Tool (LIAT), which was developed with support from the United States Agency for International Development (USAID); the Logistics Handbook (USAID); and USAID’s National Supply Chain Assessment Toolkit (NSCA), which includes the Capability Maturity Model Diagnostic Tool and National Supply Chain Key Performance Indicator Assessment (USAID). However, most of these reports and tools do not clearly distinguish or link between central, regional, and facility-level indicators. For example, no clear evidence exists that indicators developed and validated at the central and regional levels can be used for measuring SCM at facility levels. In addition, these tools cannot be readily used to understand how established facility-level SCM indicators affect and are connected to those at the regional and central levels. Furthermore, only a few studies examine associations of facility-level practice indicators to SCM.

Some facility-level indicators of practices and behaviors do exist, but many of them have not been shown to be associated with SCM, and no system is in place to measure and monitor them in a systematic way. In this study, we aimed to identify facility-level practices and behaviors and their associations to ARV SCM and to propose indicators to measure these practices and behaviors in future studies.

This study is unique in 2 ways. As previously stated, few studies have tried to expand understanding of facility-level practices and behaviors that could affect ARV SCM. Second, very limited publications or reports utilize a qualitative methodological framework of analysis to elucidate what is currently happening at the facility level as a means of determining the practices and behaviors that may have an impact on facility-level SCM.

## METHODS

We used a mixed research methodology that incorporated both quantitative and qualitative methods. The methodology included 4 phases consisting of a literature review, country and health facility selection and survey instrument development, in-country data collection, and data synthesis and analysis.

### Literature Review

First, we conducted an in-depth review of published and gray literature to identify facility-level SCM indicators that can be used to measure and categorize facility-level practices and behaviors. The literature review was performed to identify facility-level practices and behaviors with evidence for an impact on SCM, with a specific focus on availability of and access to HIV/AIDS commodities. We focused mainly on literature from low- and middle-income countries, but we also searched for management literature in business journals, which included sources from higher-income countries. Key terms used in the search were HIV/AIDS, medicines and/or pharmaceutical supply chain management, supply chain performance, supply chain performance measurement, qualitative HIV/AIDS studies, behaviors and/or practices that impact supply chain, HIV/AIDS supply chain guidelines, and terms related to each step of the supply chain (e.g., product selection, ...
forecasting, procurement, warehousing, inventory management, transportation, prescribing, and dispensing).

The review began with a search of published materials from 1990 to 2015 in the following databases: PubMed, Web of Science, and Google Scholar. We reviewed official reports published by nonprofit organizations, development agencies, and their implementing partners, in particular, USAID, MSH, John Snow Inc., and the World Health Organization. The following tools and reports were also analyzed: Managing Access to Medicines and Health Technologies (MDS-3, MSH), the Logistics Indicators Assessment Tool (USAID), Logistics Handbook (USAID), the 2013 toolkit documents from USAID’s NSCA: Capability Maturity Model Diagnostic Tool and National Supply Chain Key Performance Indicator Assessment.

To identify existing facility-level SCM indicators, we reviewed indicators related to product selection, forecasting and supply planning, procurement, warehousing and inventory management, transportation, dispensing, waste management, laboratory issuing, information management, infrastructure, human resources, demand factors, behaviors and practices, and perceptions related to HIV/AIDS treatment. In addition to facility-level indicators, supply chain function areas were also reviewed in the literature. However, after analyzing numerous supply chain tools and reports, we chose to extract and adapt the supply chain function areas from USAID’s 2013 NSCA Toolkit to categorize the facility-level practices and behaviors identified in the literature review. The main rationale for this decision was that the majority of the supply chain indicators and function areas proposed in that toolkit originated from and consolidated various aspects of other major tools.

The final 7 supply chain function areas used to classify facility-level behaviors and practices related to ARV SCM were forecasting and quantification; warehousing and inventory management; prescribing and dispensing; communication; information management; infrastructure; and human resources.

The Figure shows the theoretical framework that guided the in-depth literature review, the methods, and the analysis. The blue boxes highlight areas included in the scope of the study. The framework shows how the study focused on identifying facility-level practices and behaviors, within facility-level supply chain function areas, and then linked them with measurable indicators. This framework also demonstrates how facility-
level practices and behaviors connect to facility- and central-level SCM outcomes. We also considered health system design factors such as human resources, organization, and leadership. The orange boxes are included to capture the entire system, but they are not part of this study.

**Country and Health Facility Selection and Survey Instrument Development**

Country selection was carried out in coordination with the USAID-funded Systems for Improved Access to Pharmaceuticals and Services (SIAPS) Program, implemented by MSH. Countries selected were PEPFAR-funded African countries based on the source of funding for this research study.

The additional selection criteria pertained to language spoken (at least 1 francophone and 1 anglophone country), geographical variation (west/south/east), and availability and preparedness of country teams. Based on these criteria, 1 francophone and 2 anglophone countries were selected that also represented varied geographic areas: Anglophone/West/Southern Africa (Namibia), Anglophone/South/Eastern Africa (Swaziland), and Francophone/West/Central Africa (Cameroon). Further information was gathered based on each country selected, including a desktop review of the official, published, and gray literature on the ARV supply chain in each country.

We collaborated with local country experts from the MOH, particularly managers of national HIV/AIDS programs and SIAPS colleagues, to select the health facilities included in the study. Country experts assisted the study team in obtaining further reports and data that were not available publicly or online. Due to the variation and different types of lower-level facilities (health posts, clinics, etc.) in all 3 countries, hospitals were chosen as the facility unit of analysis due to more standard patient care responsibilities, availability of pharmacies, and types of health care workers and human resources. With technical assistance from national HIV/AIDS program managers, 4 hospitals in different geographic areas from each country were identified, for a total sample size of 12 hospitals.

The hospitals included in the study were selected based on indicators that we reviewed with each collaborating country study team. The indicators used to select the hospitals (listed below) were consistent within a country, but varied across the 3 countries. This variation was due to the 3 countries having different health information systems, which led to facilities across countries having different comparable and available indicators. The following 5 supply chain–related indicators were chosen to select facilities in Namibia:28,29

- Storage condition: cleanliness, tidiness, appropriate arrangement of pharmaceuticals, and temperature monitoring practices
- Inventory management and quantification: stock card use, cold chain management, and interim orders
- Use of the Electronic Dispensing Tool for stock management (such tools help maintain basic patient profile information, medicine history, and other data that are essential for the dispenser to know at the time of dispensing)
- Completeness of ART reports submitted to the Ministry of Health and Social Services (MOHSS)
- Availability of adequate ART stock on hand

The following 2 supply chain–related indicators were used in Cameroon:

- Discrepancies between patient figures and stock records
- Discrepancies between physical stock count and stock records

The following supply chain–related indicator was used in Swaziland:

- Months of stock on hand

Based on the in-depth literature review as well as country protocols for the management of HIV/AIDS patients in the participating countries, we developed qualitative interview guides to identify the practices and behaviors at the facility level that could have an impact on ARV SCM. Facility-level practices and behaviors were included in the guide for the following activities: prescribing, counseling, dispensing, clinical services, data-generating activities, data analysis, data reporting, forecasting, ordering, handling of expired stock, behaviors if a stock-out happens, submission of reports and orders, placing emergency orders, communication within and between facilities, and more general practices and behaviors regarding management leadership style and the quality of communication between facilities and higher-level SCM. All these activities fell within the 7 supply chain function areas.

Interview guides were developed to facilitate focus group discussions with key personnel at the facility level and interviews with key informants.
at central health system levels (e.g., the MOH, HIV/AIDS National Program, Central Medical Stores [CMSs]). We identified relevant interviewees based on their reported assigned tasks and daily responsibilities in each of the health facilities in the provision of ARVs and HIV services. We did not rely solely on using job titles or job descriptions because they do not necessarily correspond to actual day-to-day responsibilities and were not consistent across facilities.

In-Country Data Collection
One-week country visits occurred in January and February of 2014. Two of the authors, both from academia, conducted the interviews solely for this study in order to reduce any bias from donor-related activities, assessments, or evaluations.

First, key informant interviews were conducted with broader health system representatives from the MOH, HIV/AIDS National Program and regional offices, CMSs, and the World Health Organization. Second, key personnel focus group interviews were conducted with health care workers, pharmacists, and managers who interact with patients with HIV/AIDS arriving to the facility (hospital) for HIV medications. No key personnel were key informants.

Data Synthesis and Analysis
Interviews were recorded through interviewers’ notes and later transcribed. A rubric of all practices and behaviors was developed to consistently identify themes and outliers in each of the key practices and behaviors within and across each country. A qualitative thematic analysis was performed to identify patterns in the practices and behaviors at the facility level that could have an impact on ARV SCM. A hybrid of deductive and inductive coding was used; meaning information identified from the literature at the beginning of the study was used as a starting point for analysis (deductive), and new themes and codes later emerged from the country-level data analysis of the interviews (inductive). All facility-level practices and behaviors observed in the countries were mapped to the 7 supply chain function areas adapted from the USAID’s NSCA Toolkit.

Due to the extensive amount of qualitative results on all 30 practices and behaviors analyzed, we selected and summarized qualitative information for 7 of them in the results. We selected these 7 because they provide clear examples of how behavior and practice vary between and within countries across the supply chain function areas. In addition, the in-depth literature review and individual country results revealed that these 7 practices and behaviors had an association with SCM.

The same 2 authors who carried out the in-country interviews and focus groups also carried out the data synthesis and analysis. Data synthesis and analyses were reviewed independently by authors 1 and 4 using a rubric of all identified practices and behaviors. For the small number of cases of disagreements, authors 1 and 4 discussed and resolved the discrepancy together. If the disagreement could not be resolved, author 2 was brought in and resolution was achieved by majority.

All major SCM measurement tools and relevant literature were then reviewed for a final time to identify indicators that mapped to each of the observed facility-level practices and behaviors. All indicators that were identified for each facility-level behavior and practice were listed. For those with no identifiable indicator, we used the results of the in-country interviews to propose new indicators. It should be noted that while this study used ARVs to understand gaps in practices at the facility level, many of the findings can be applied more broadly to medicines in an integrated setting.

Country Context
Table 1 summarizes relevant information on country context for Cameroon, Namibia, and Swaziland with respect to spending on HIV/AIDS, HIV prevalence, and demographics. The Namibian government has made the largest investment in HIV/AIDS spending, with 64% of the total HIV/AIDS budget coming from the public sector in 2014. Swaziland covers 38% of HIV/AIDS program costs with government funds, and Cameroon spends 15%. Consequently, Cameroon had the largest foreign donor investment in HIV/AIDS in 2014 (70%–80% of HIV/AIDS funding), with most of it coming from the Global Fund. Of the 3 countries, Cameroon has the lowest HIV prevalence, at 4.8% of 15–49-year-olds. Swaziland has one of the highest global HIV prevalence rates, with 26% of 15–49-year-olds living with HIV.

Namibia, the largest country in the study, has a population of only 2.3 million people, making it one of the least populated countries in the world. Namibia not only has a shortage of health care workers but also a shortage of supply chain personnel to work in the public health supply chain. The MOHSS oversees 14 regions with 350 public health facilities and manages the CMS.
distribution systems in Windhoek, Rundu, and Oshakati (the latter 2 being multiregional medical depots [MRMDs]). The distribution of medicines is carried out by CMS Windhoek to public health facilities in nearby central and southern Namibia and to the distant MRMDs of Rundu and Oshakati to cover a wide swath of Namibia’s territory. MRMDs and public health facilities deliver medicines to health centers, clinics, and hospitals. Namibia operates an integrated pharmaceutical supply chain; ARVs do not have a parallel system of delivery, thus they are distributed with all other medicines from CMSs to clinic levels.

Cameroon is about half the size of Namibia but its population is tenfold higher (23.4 million inhabitants). The National AIDS Commission manages Cameroon’s HIV/AIDS program. The national program manager is responsible for all programmatic aspects related to quantification and forecasting of ARVs and implementation of HIV/AIDS policies, among other duties. Cameroon has 10 regional medical stores (RMSs; centres d’approvisionnement pharmaceutiques des regionaux [CAPRs]) in each of its 10 administrative regions. CAPRs deliver medicines to health facilities and all receive ARV stock procured by Cameroon’s CMS based in Yaounde (Centrale Nationale d’Approvisionnement en Médicaments Essentiaux et Consommables Médicaux). CAPRs distribute ARVs to the 124 ART sites based on health clinic needs; however, when ARVs are only available in limited quantities or facing stock-outs at the national level, ARVs are rationed to health facilities.

Swaziland is a small country with a population of 1.4 million, making it the most densely populated country in this study. Swaziland has an integrated pharmaceutical supply chain for medicines and medical supplies. CMSs are responsible for the SCM of all health commodities including medicines in the public sector. ARVs are directly distributed from CMSs on a monthly schedule to 45 hospitals and health facilities. ARVs are further delivered down the supply chain from hospitals and health facilities to what are known as baby (feeder) clinics.

**Ethical Approval**

The study was approved by the Institutional Review Board (IRB) of the Harvard T.H. Chan School of Public Health (protocol number IRB13-3167) and was determined to be exempt from obtaining written informed consent. The IRB determined that the protocol meets the criteria for exemption per the regulations found at 45 CFR 46.101(b)(2) and any additional review by the IRB was not required. In country, no local IRB approval was required; however, the research team obtained written authorization to conduct the study from the MOH of Cameroon, Namibia, and Swaziland. All participants in this study were adults (≥18 years). The purpose of the study was explained to participants, and verbal informed consent was obtained from all individuals participating in individual and focus group qualitative, semistructured interviews.

**RESULTS**

In total, 52 semistructured key informant and personnel interviews were conducted across the 12 hospitals included in the study. The number of
key informants and personnel that were interviewed in each hospital varied depending on how many individuals were available on that particular day. For example, in Namibia across the 4 health facilities visited, 8 key informant/personnel interviews were conducted for an average of 2 individuals in each hospital. There were more individuals available for key informant interviews at the broader health system level in Namibia and Cameroon than in Swaziland. Table 2 details the number of key informant interviews conducted in each country and the type of interview.

Table 3 shows the 30 facility-level practices and behaviors that were identified through key informant and personnel interviews and mapped to 7 facility-level supply chain function areas. Some of these practices and behaviors are activities that are closely linked with routine activities for that particular supply chain function area and the literature supports a link between better completion of this activity and improved SCM. For example, calculation of minimum and maximum (min-max) buffer stock is an activity that is practiced widely across many facilities as part of forecasting and quantification exercises, although specific types of calculation differ. The MOH and donors often provide trainings on the different calculations that can be used to verify and document the maximum and minimum number of months of stock that should remain on the shelf, which should lead to lower stock-outs.7

Other identified practices and behaviors are not routinely carried out as part of that function area, and the literature contains minimal evidence on the link between them and any type of improved SCM. For example, communication between pharmacy team members and the clinical staff dispensing the medications is routinely practiced to different degrees of effectiveness across all the facilities included in the study. However, the literature includes little documentation on these types of communications and the relationship of communication with SCM.

In order to shorten the length of the full qualitative results, we selected 7 of the 30 practices and behaviors associated with ARV SCM as clear examples of the variation in supply chain function areas within each country (Table 4). These practices and behaviors are order verification before submission to the central/regional level; actions taken when stocks received from CMS/RMS; communication with higher-level SCM; communication with affiliated facilities; change in prescription during stock-out; change in dispensing during stock-out; and actions to ensure patient adherence. Additional qualitative information on all 30 practices and behaviors can be found in the SIAPS report Facility Level Practices and Behaviors That Affect the Performance of the Supply Chain of Antiretroviral Medicines.7 The 7 practices and behaviors reported and discussed in the current study describe what health care workers actually do (what they describe doing) and whether or not it is in line with guidelines for receiving a shipment, making an order, communicating among staff, and managing medicines and commodities.

**Order Verification Before Submission to the Central/Regional Level (Supply Chain Area: Forecasting and Quantification)**

Order verification refers to a health facility staff member rechecking the ARV requisition request (verifying calculation, order, and inventory stock) before an order is sent to the central or regional level. We noted variation across facilities within each of the 3 countries with respect to how orders are verified. In Namibia, 2 facilities have a regional pharmacist and/or senior pharmacists who verify

<table>
<thead>
<tr>
<th>TABLE 2. Semistructured Key Informant and Personnel Focus Group Interviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interviews</td>
</tr>
<tr>
<td>Broader health systems interviews</td>
</tr>
<tr>
<td>Hospital interviews (staff involved in HIV/AIDS services and ARV management)</td>
</tr>
<tr>
<td>Total interviews (N=52)</td>
</tr>
<tr>
<td>Total facilities in the study (N=12)</td>
</tr>
</tbody>
</table>

Abbreviation: ARV, antiretroviral.
<table>
<thead>
<tr>
<th>Facility-Level Supply Chain Function Area and Practice/Behavior</th>
<th>Description of Best Practices From the Literature and Observed Practices From This Study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Forecasting and Quantification</strong></td>
<td></td>
</tr>
<tr>
<td>1. Calculation of minimum and maximum buffer stock</td>
<td>Calculation of the minimum and maximum levels of pharmaceutical stock needed over a specified time period, taking into consideration buffer stock, stock used during lead time, and order quantity for one supply period.5</td>
</tr>
<tr>
<td>2. Use of electronic systems</td>
<td>Use of electronic systems assist in the tracking of services and products delivered to patients. Furthermore, such systems also help to fulfill new monthly orders and maintain stock records, while also assisting in reporting such records to higher-level offices.</td>
</tr>
<tr>
<td>3. Use of national guidelines as reference for estimation of needs and reporting</td>
<td>Use of guidelines in inventory control improves poorer performance of logistic systems.38</td>
</tr>
<tr>
<td>4. Order verification before submission to the central/regional levela</td>
<td>A health facility staff member rechecking the ARV requisition request (verifying calculation, order, and inventory stock) before an order is sent to the central or regional level leads to fewer order verification errors.</td>
</tr>
<tr>
<td>5. Order fill rate calculation</td>
<td>Order fill rate should be calculated to cut down number of emergency and/or unfilled orders.</td>
</tr>
<tr>
<td>6. Late ordering of medicines</td>
<td>Staff should be consistently aware of order dates and treat them as a potential problem so as to avoid late orders</td>
</tr>
<tr>
<td>7. Frequency of issuing emergency orders</td>
<td>A study in Mali found that emergency orders of stock are required as facilities receive only about 25% of what they request.39 Emergency orders were not reported as frequent or an issue.</td>
</tr>
<tr>
<td><strong>Warehousing and Inventory Management</strong></td>
<td></td>
</tr>
<tr>
<td>8. Actions taken when stock received from CMS/RMSa</td>
<td>Any newly received or issued products are recorded in stock-keeping records. Entries are further updated either when stock is counted during a physical inventory, or as soon as a loss is noticed.5</td>
</tr>
<tr>
<td>9. Control of access to stock</td>
<td>Security, monitoring, and auditing are some of the methods to prevent stock-outs and losses.26</td>
</tr>
<tr>
<td>10. Decision on whether to redistribute short-dated stock</td>
<td>Redistribution of short-dated stocks increases the complexity of the supply chain and miscommunication of stock levels between facility and central levels.</td>
</tr>
<tr>
<td>11. Location and condition of storage (whether all in one place or separate rooms)</td>
<td>Good inventory control includes appropriate storage space, stock rotation, stock arrangement, cleanliness, security, and fire prevention.26</td>
</tr>
<tr>
<td>12. ARVs stored separately from other medicines</td>
<td>Due to funding requirements, many ARVs are stored in separate storage areas from other medicines. Access of staff to ARVs is limited as well to prevent theft and diversion.</td>
</tr>
<tr>
<td>13. Assigning responsibility of inventory management tasks</td>
<td>In most facilities, a trained nurse, pharmacy assistant, or pharmacist is assigned to manage ARVs. In some facilities a schedule and description of tasks for staff is available and implemented.</td>
</tr>
<tr>
<td>14. Frequency of balancing stocks (checking stock cards vs. physical count)</td>
<td>Stock status of each product in storeroom should be assessed regularly (monthly) by staff, comparing the quantities on hand with the quantities that have been entered in inventory control cards.5</td>
</tr>
<tr>
<td><strong>Prescribing and Dispensing</strong></td>
<td></td>
</tr>
<tr>
<td>15. Change in ARV prescription during stock-outa</td>
<td>SOPs are needed for the prescribing process in the event of stock-outs to standardize actions among prescribers.</td>
</tr>
<tr>
<td>16. Change in dispensing of ARVs during stock-outa</td>
<td>Written SOPs are recommended to improve consistency and quality of the dispensing process.26 SOPs are needed to standardize actions during the dispensing process.</td>
</tr>
</tbody>
</table>

Continued
### TABLE 3. Continued

<table>
<thead>
<tr>
<th>Facility-Level Supply Chain Function Area and Practice/Behavior</th>
<th>Description of Best Practices From the Literature and Observed Practices From This Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>17. Actions to ensure patient adherence (e.g., pill count)⁴⁰</td>
<td>SOPs are needed for monitoring adherence (e.g., whether to perform pill counting) to ARVs.</td>
</tr>
<tr>
<td><strong>Communication</strong></td>
<td></td>
</tr>
<tr>
<td>18. Communication within the pharmacy team</td>
<td>A positive team dynamic can be achieved via regularly scheduled weekly/biweekly internal meetings.</td>
</tr>
<tr>
<td>19. Communication within the facility</td>
<td>Active communication between pharmaceutical and nonpharmaceutical staff regarding shortages and stock-outs is recommended to increase consistency and accurate recording of prescriptions.</td>
</tr>
<tr>
<td>20. Communication with higher-level supply chain management⁴⁰</td>
<td>Improved facility-level SCM performance can be achieved more easily via robust relationships with the regional and central personnel.</td>
</tr>
<tr>
<td>21. Communication with affiliated facilities⁴</td>
<td>An increased in accurate reporting and forecasting at the main facility is a potential byproduct of positive and regular communication with any and all affiliated facilities.</td>
</tr>
<tr>
<td>22. Communication with hospital executives</td>
<td>Key informants report that direct lines of communication between pharmaceutical staff and hospital executives is recommended to address and avoid shortages and stock-outs.</td>
</tr>
<tr>
<td><strong>Information Management</strong></td>
<td></td>
</tr>
<tr>
<td>23. Interaction between clinical and dispensing/stock systems</td>
<td>Most facilities do not have linkage between clinical and dispensing information systems. Swaziland does have linked systems and key informants report frequent backlogs on prescription input.</td>
</tr>
<tr>
<td><strong>Infrastructure</strong></td>
<td></td>
</tr>
<tr>
<td>24. ARV clinic/pharmacy separate from main pharmacy</td>
<td>ARV clinic/pharmacy was observed to be separate from the main pharmacy in some facilities and integrated with others.</td>
</tr>
<tr>
<td><strong>Human Resources</strong></td>
<td></td>
</tr>
<tr>
<td>25. Training on stock management</td>
<td>An individual’s technical ability, personality, and position within the supply chain had a significant impact on supply chain performance.⁴⁰</td>
</tr>
<tr>
<td>26. Leadership/management style of the pharmacy</td>
<td>Key informants reported multiple leadership/management styles of the pharmacies. Some were managed/led by regional and senior level pharmacists, others by pharmacists, pharmacist assistant physicians or nurses. Consistent management organization and leadership across pharmacies can improve supply chain performance.</td>
</tr>
<tr>
<td>27. Leadership management style of the clinic</td>
<td>Key informants reported that clinics were managed/led by physicians who attend HIV patients and other patients.</td>
</tr>
<tr>
<td>28. Attitude to workload of pharmacy staff</td>
<td>Pharmacist assistants and nurses in some facilities reported that workloads were too high, leading to unfinished daily activities, including those linked to supply chain management.</td>
</tr>
<tr>
<td>29. Guidelines for providers in the event of a stock-out</td>
<td>There are no standardized guidelines for providers for what to do in the event of a stock-out.</td>
</tr>
<tr>
<td>30. Implementation of policies on prescribing and dispensing</td>
<td>Some key informants reported having clear policies of not allowing patients to leave without any medicines.</td>
</tr>
</tbody>
</table>

Abbreviations: ARV, antiretroviral drug; CMS, central medical store; RMS, regional medical store; SCM, supply chain management; SOP, standard operating procedure.

⁴ These practices and behaviors are associated with SCM more than others and are described in detail in the results section.
<table>
<thead>
<tr>
<th>Facility-Level Supply Chain Function Area</th>
<th>Practice/Behavior</th>
<th>Sample Measurable Indicators From Existing Research and Literature</th>
<th>Potential New Facility-Level Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forecasting and quantification</td>
<td>Order verification before submission to the central/ regional level</td>
<td>• Formal work plan and/or schedule for quantification^2^6</td>
<td>• Orders are verified by staff prior to sending</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Average order entry time and order entry accuracy^9</td>
<td>• Second-stage order verification by staff member other than the person who filled the order</td>
</tr>
<tr>
<td>Warehousing and inventory management</td>
<td>Actions taken when stocks received from CMS/RMS</td>
<td>• Average put-away accuracy and put-away time^9</td>
<td>• Immediate shelving of stock upon arrival by appropriate staff member</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Verification of stock arrival and shelving procedures</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Reported discrepancies between what was in the order placed and what was actually received</td>
</tr>
<tr>
<td>Prescribing and dispensing</td>
<td>Change in prescription</td>
<td>• Perception of physician—if physicians are perceived to be professionally competent, pharmacy staff may model their behavior on physician prescribing patterns. Presence of some medical malpractice could also influence the pharmacy staff’s behavior.</td>
<td>• Standardized procedure/formal communication among prescribers to adjust prescriptions during stock-outs followed</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Number of patients switched to another regimen due to stock-outs, then switched back to the old regimen or kept on the new regimen when the drug becomes available</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Changes in prescriptions recorded at the pharmacy</td>
</tr>
<tr>
<td></td>
<td>Change in dispensing during stock-out</td>
<td>• No existing indicators</td>
<td>• Standardized procedure/formal communication among pharmacy staff regarding the amount to dispense during stock-outs followed</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Changes in dispensing recorded</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Discrepancies in what was prescribed and dispensed recorded</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Procedure established/followed when one of the medicines in a regimen is stocked out, and what happens to the other medicines (e.g., thrown out, given to someone else)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Number of stock-outs for pediatric formulation affecting management of adult ARV stocks</td>
</tr>
<tr>
<td>Action to ensure patient adherence</td>
<td>• Regular pill counting^2^6</td>
<td>• Percentage of patients with full adherence to ART (i.e., no doses missed in the 3-day recall period)</td>
<td>• Pill counting conducted</td>
</tr>
<tr>
<td></td>
<td>• Percentage of patients with full adherence to ART (i.e., no doses missed in the 3-day recall period)</td>
<td>• Changes in dispensed medicines recorded</td>
<td></td>
</tr>
</tbody>
</table>

Continued
the orders. For example, at 1 hospital, the regional pharmacist always verifies stock card records on visits, even if the senior pharmacist has verified the order and inventory check of the pharmacy assistant. In the other 2 facilities visited in Namibia, no one checks the order before it is sent. In Swaziland, 3 of the 4 hospitals have a larger pharmacy team in which a senior pharmacist supervises work. This management structure enabled verification and review of orders before submission. In Cameroon, all 4 hospitals have their orders verified by the hospital coordinator.

**TABLE 4. Continued**

<table>
<thead>
<tr>
<th>Facility-Level Supply Chain Function Area</th>
<th>Practice/Behavior</th>
<th>Sample Measurable Indicators From Existing Research and Literature</th>
<th>Potential New Facility-Level Indicators</th>
</tr>
</thead>
</table>
| Communication                            | Communication with higher-level supply chain management | • Average percentage of days covered by ARVs dispensed for a sample of patients for a defined period (180 days)  
• Percentage of patients who experienced a gap in ARV availability of more than 30 days in a row during the same defined period  
• Percentage of patients who attend on or before the day of their appointment  
• Percentage of patients who come within 3 days of their appointment | • Perception of relationships between ARV manager/coordinator and regional offices  
• Frequency of communication (times/month, times/year)  
• Number of times the regional office “checks” on each pharmacy (times/month, times/year)  
• Perception of relationships between pharmacy and central medical store  
• Perception of support and good supervision the ARV manager thinks they receive from the regional pharmacist |
| Communication                            | Communication with affiliated facilities | • No existing indicators | • Type of communications that occur between the facility and its affiliated facilities (e.g., outreach sites, baby clinics)  
• Procedure for affiliated facilities placing orders followed  
• Frequency that affiliated facilities place orders with the higher-level facility (times/month, times/year) |

Abbreviations: ARV, antiretroviral drug; ART, antiretroviral therapy; CMS, central medical store; RMS, regional medical store.
Actions Taken When Stocks Are Received From CMS/RMS (Supply Chain Area: Warehousing and Inventory Management)

Actions taken when ARVs are received from the CMS/RMS refers to how and when new medications are received and added to the stock-keeping record. When stock arrives, some facilities quickly and diligently take actions, while others delay checking the physical count of medicines, opting instead to defer recording the receipt of stock and postpone storage.

In Namibia, 1 hospital prioritizes the counting, storing, and recording of new stocks upon arrival. The 3 other facilities experience delays; actions taken are less urgent or they are performed as staff becomes available. In Cameroon, 3 of 4 facilities take action when stocks are received. In 1 of these facilities, a staff member travels to the regional CAPRs to collect stocks, check physical count, and match them with the invoices. When this staff member returns to the hospital, stock cards are updated immediately with expiration dates and quantity. In 2 other facilities in Cameroon, staff members travel to the CAPRs to collect stocks and check physical counts to ensure stock cards are updated immediately upon stock arrival. In Swaziland, 2 of the 4 hospitals unpacked the new stocks when they “have time,” which can sometimes take days. The other 2 facilities unload and unpack the stocks as soon as they arrive.

Communication With Higher-Level SCM (Supply Chain Area: Communication)

Communication by the health facilities with their regional offices refers to the dialogue and discussion around supply chain issues between individuals in the health facilities and higher-level authorities who might be MOH personnel or individuals at the CMS. In Namibia, communication with higher-level regional management in 2 facilities is reported to be quite frequent (daily communication with the regional pharmacist). The other 2 facilities in Namibia report poor communication in that the regional pharmacist “never” visits the facility.

In Cameroon, 1 facility contacts the CAPR through the hospital coordinator; in another facility, the nurse directly contacts the regional office whenever a shortage occurs. The other 2 hospitals both call their respective CAPRs on a monthly basis to check their stock levels prior to placing an order for ARVs.

Communication with higher-level SCM does not take place in Swaziland because there is no regional pharmacist; however, key informants indicated in interviews that having a regional pharmacist would be helpful. Communication with CMS is done on a frequent basis by all facilities, but the CMS is consistently slow to respond to facility requests. Therefore, health facilities at times use other sources to obtain the medicines they need, such as reaching out to donors or other health facilities to inquire about stock and delivery availability.

Communication With Affiliated Facilities (Supply Chain Area: Communication)

Communication with the affiliated facilities refers to the dialogue and discussion of supply chain issues with other ARV clinics, other health posts, or other health personnel such as the physician who comes for outreach visits. Most facilities in Namibia reported the communication with their affiliated facilities is strong, which assists in maintaining appropriate stock levels and orders at the affiliated facilities. In Swaziland, 2 facilities reported having positive communication with their affiliated clinics, while 2 other facilities expressed concern with their clinics due to a backlog of data and nonsubmission of orders on a timely and necessary basis. No trends were reported for Cameroon because there are no affiliated facilities.

Change in Prescription and Dispensing During Stock-Out and Action to Ensure Patient Adherence (Supply Chain Function Area: Prescribing and Dispensing)

The results highlight 3 key themes in prescribing and dispensing that are noteworthy: change in ARV prescription during stock-out, change in dispensing of ARVs during stock-out, and actions to ensure patient adherence (e.g., pill count). This section summarizes some of the trends within each country with regard to changing a patient’s ARV regimen and changing dispensing practices during periods of stock-out.

In all 3 countries, each of these prescribing and dispensing practices frequently occurs in all facilities. For example, 1 facility reports changing the ARV formulation for children if there is a stock-out of a particular formulation. If the child is currently using an ARV in tablet form but the facility has a stock-out, the physician will prescribe the syrup formula instead. There are also reports of modifying the duration of the prescription. For example, if a facility is short-stocked on an ARV medication and the physician is planning to travel to reach clinics located at a further distance, the
physician prescribes only 1 month, instead of the regular 2 months of ARVs. The patient must then return to the facility in 1 month, which could potentially have a negative impact on adherence. Facilities also reported changing patients to different first-line or second-line regimens that are available when a shortage occurs. However, the patient often does not resume the original regimen when the ARV is back in stock and available. In summary, patients continue with the new first- or second-line regimen unless clinical side effects are present. These regimen changes can affect consumption patterns, forecasting, and procurement and significantly contribute to consumption pattern oscillations and a lack of predictability.

Facility-Level Practices and Behaviors With Existing and New Indicators

Existing literature and tools contain a number of indicators for forecasting and quantification and warehousing and storage. However, fewer standardized indicators to measure how facility-level behaviors and practices can affect SCM exist for prescribing and dispensing practices and communication with internal and external teams. Table 4 summarizes the mapping exercise for 7 of the 30 practices and behaviors associated with ARV SCM that were described previously, and maps each of these practices and behaviors to existing measurable indicators from the literature. For those without measurable indicators, new indicators based on country results are suggested that can potentially measure facility-level practices and behaviors and their impact on ARV SCM. Furthermore, in cases in which the qualitative data yielded important findings, new indicators are also proposed for behaviors and practices that already had some prior measures. The Supplement summarizes the mapping for all 30 practices and behaviors.

DISCUSSION

The results of this study are derived from focus group discussions with key personnel at the hospitals and semi-structured interviews with key informants consisting of higher-level health system players. The results identify 30 practices and behaviors that may affect ARV SCM, focusing on several key areas that are associated with improved SCM: order verification, actions taken when ARV stock is received, changes in prescribing and dispensing during stock-out, actions to ensure patient adherence, and communication with other affiliated facilities and higher-level SCM. We developed measurable indicators for future research, focusing on practices and behaviors related to prescribing and dispensing, communication, and human resources. With a focus on ARVs, the study results provide insight into current SCM at the facility level in Cameroon, Namibia, and Swaziland, and we make recommendations based on the findings that are considered best practices.

It was reassuring that pharmacy staff in many of the hospitals where interviews were conducted routinely calculated min-max amounts of stock, but stock-outs of key ARV products were still an issue. Therefore, despite routinely performing these calculations, staff were not able to apply the results to prevent stock-outs from happening. Several factors might explain this discrepancy, including (1) a lack of understanding and ability to calculate min-max amounts of stock followed by application of the results, (2) lack of monitoring the calculations to ensure accuracy and to link the min-max amounts to the actual amount of stock on the shelf, (3) improperly filled orders, and (4) other miscellaneous reasons related to stock assessments. A number of other, broader contextual factors were present, including orders being placed but not filled. The current study focused on practices and behaviors at the facility level, and future research will need to incorporate additional measures and analysis of upstream contextual factors.

The findings also indicate that for some supply chain functions (forecasting and quantification and warehousing and inventory management), national guidelines and/or best practices need to be standardized and disseminated to all facilities. More specifically, information from the qualitative interviews and focus groups suggests that trainings using standardized guidelines can strengthen the following practices and behaviors: order verification, redistribution of stock, emergency orders, control and access to stock, location and condition of storage, and where ARVs are stored. Trainings should be available for all facilities and not dependent on their distance from a regional office. For example, the process of redistributing stock between facilities and requesting an emergency order was extremely varied and should be reviewed to determine best practices to be communicated to all facilities. Currently, these 2 behaviors frequently occur together and are not standardized, and redistribution often occurs between facilities using informal processes. Since emergency orders are often sent to the central level at the same time that redistribution is

Some supply chain functions need standardized national guidelines and/or best practices for dissemination to all facilities.

The results of this study identify 30 practices and behaviors that may affect ARV SCM.
happening at the lower level, the central level is unaware of what has been redistributed and where, which is a broader system design and reporting issue.

The impact of prescribing and dispensing practices needs to be understood clinically to determine if any of the practices and behaviors identified in this study have a negative impact on the health outcomes of the patient. There is a lack of published literature and academic understanding exactly how and to what extent these practices affect patient health. For example, changing a regimen due to a stock-out is likely to affect adherence, may cause adverse drug reactions, and could pose risks for treatment resistance. Prescribers and dispensers frequently need to manage the reality of frequent ARV stock-outs and therefore must adjust patient doses and regimens; communication between providers and facilities about these changes is minimal, and a consistent method to record this critical information is lacking. At the time of the study, Swaziland was the only country that had an electronic patient record, including prescriptions, linked with a new pharmacy electronic system. However, implementation and linking of the 2 systems had some issues, causing backlog of requests. Namibia had an electronic dispensing tool, but it was not linked with clinical records. Health facilities are often on their own in dealing with shortages and stock-outs since there are no national standard operating procedures (SOPs) or standardized prescribing guidelines specific to these emergency issues.

While the importance of communication has been identified as a key aspect of appropriate management in health care in general, little research has investigated how better communication could be linked to SCM. Five different behaviors are linked to communications, including communication within the pharmacy team; communication within the facility; communication with higher-level hospital/clinic executives; communication with higher-level supply chain (regional and central medical stores); and communication with affiliated facilities. Communication within the facility, communication with affiliated facilities, and communication with higher-level SCM offices were newly identified practices and behaviors with potential indicators to be field tested.

**Recommendations**

Based on our findings, we propose recommendations to improve SCM in 4 main facility-level areas.

- **Understanding and monitoring of routine min-max calculations.** Pharmacy staff should be monitored to perform the min-max calculations. This monitoring can happen within the facility by another colleague, such as a hospital coordinator, or by a higher-level regional pharmacist. Many hospitals relied on their regional pharmacist for periodic visits to check on specific practices and behaviors, reporting that the monitoring improved their SCM. Facilities in regions that did not have a regional pharmacist reported that having one would be helpful in such monitoring activities.

- **Development of national guidelines and norms.** Best practices and guidelines should be developed and monitored at the facility level for order verification, redistribution of stock, emergency orders, control and access to stock, location and condition of storage, and where ARVs are stored.

- **Understanding the impact of prescribing and dispensing practices on health outcomes.** Prescribers and dispensing staff would certainly benefit from guidelines, SOPs, and trainings on ARV regimen switches and managing shortages. Additional research is needed to understand the following: how prescribing and dispensing practices change during shortages and stock-outs; the frequency of regimen changes practiced in the country; and the impact of regimen changes on health outcomes including adherence and resistance.

- **The importance of communication.** Communication has been more frequently measured and studied in hospital management, business, and organizational fields; however, how different types of communication can affect SCM has not been thoroughly examined or measured with standardized indicators. This gap holds true at the facility, regional, and central levels.

Our recommendations are most likely applicable to other countries around the world that are experiencing similar difficulties with their own ARV supply chains. Further research, particularly field-based, pilot testing of the new, potential indicators for communication, human resources, and prescribing and dispensing identified in this study, is needed to understand their impact on SCM. Once these potential indicators are piloted for validity and reliability, causal study designs can be developed to test which practices and behaviors have a direct, measurable impact on facility-level SCM.
Limitations

Our study yielded some interesting findings, but it had some limitations. First, it was limited to selecting PEPFAR-supported countries based on availability and preparedness of the country teams. Secondly, the facility selection indicators varied across all 3 countries using information from country experts and available in-country data. Due to this wide variation in selection criteria for facilities across the countries, cross-country or intracountry comparisons or causal conclusions could not be made. Consequently, for each generalized result given above, a country-specific example is provided. While this is a major limitation, which is exacerbated by the sample size, it provides a baseline and unique opportunity to investigate these associations in more detail within each country. Furthermore, our study did not consider how gender, ethnicity, socioeconomic status, or other power dynamics may influence the lens through which the key informants viewed the issues. Finally, the relationship between facility- and central-level indicators (i.e., CMSS) is not explored in this article; however, improvements in facility-level practices and behaviors should logically lead to better SCM at the central level.

Due to these limitations, no cross-country or intra-country comparisons or conclusions could be made about potential linkages between the facility-level practices and behaviors identified and SCM outcomes.

CONCLUSIONS

This study identifies gaps in the measurement of facility-level practices and behaviors in the ARV supply chain. The results identify 30 facility-level practices and behaviors associated with SCM; 15 of the 30 do not have any existing indicators for measurement from literature and research. Potential new indicators are proposed based upon the analysis of the in-country qualitative results.

Improved practices and behaviors are recommended for implementation, such as offering trainings, developing SOPs and national guidelines related to communication with internal and external staff, developing new guidelines on how to adjust prescribing and dispensing patterns during shortages or stock-outs, and limiting access to ARVs in storage to selected staff at the facility. The study results provide insight into the way that ARV supply chains are currently managed at the facility level in 3 countries of sub-Saharan Africa. The results may serve as the basis for additional research, both within and across countries, to examine how improving specific behavior and practices affects SCM.

Acknowledgments: We would like to thank all the participants who provided information during the interviews for this study. We would also like to thank Helena Walkowiak and Ruth Musila from SIAPS for their contributions as technical reviewers of this article. We also recognize the approvals and support we received from the Ministries of Health and USAID missions in Cameroon, Namibia, and Swaziland. This article is based on our unpublished report titled, “Facility-Level Practices and Behaviors That Affect Performance of the Antiretroviral Medicine Supply Chain” (http://siapsprogram.org/publication/facility-level-practices-and-behaviors-that-affect-performance-of-the-antiretroviral-medicine-supply-chain/), which is referenced throughout the current article.

Funding: This article was made possible by the generous support of the American people through USAID, under the terms of cooperative agreement number AID-OAA-A-11-00021. The authors led the study design, data collection, analysis, and interpretation of the data as well as the writing of the manuscript. The funding entities, USAID and MSH, provided input into the study design and overall research concept.

Disclaimer: The contents of this article are the responsibility of MSH and do not necessarily reflect the views of USAID or the United States Government.

Competing Interests: None declared.

REFERENCES

Associations Between Health Facility Practices and ARV Supply Chain Management


Peer Reviewed

Received: February 6, 2019; Accepted: April 30, 2019

Cite this article as: Bowser D, Krech L, Mabirizi D, Chang AY, Kapaon D, Bossert T. Associations between practices and behaviors at the health facility level and supply chain management for antiretrovirals: evidence from Cameroon, Namibia, and Swaziland. Glob Health Sci Pract. 2019;7(2):300-316. https://doi.org/10.9745/GHSP-D-19-00063

© Bowser 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-19-00063
Planning for Outcomes (P4O) Modeling Tool: Estimating the Impact of Changing the Proportion of Injectable Progestins in the Contraceptive Method Mix

Elena Lebetkin, a Xiaoming Gao, a Douglas Taylor, a Lauren Y. Maldonado, a Abdulmumin Saad, b Markus J. Steiner, a Laneta J. Dorflinger, a Kavita Nanda, a Timothy D. Mastro a

The interactive deterministic online modeling tool P4O allows users to estimate how changing the proportion of injectable progestins in the contraceptive method mix might affect HIV and maternal and child health outcomes. With careful consideration for women’s individual choices, policy makers and program planners may use country-specific results to help inform programming and policy decisions.

ABSTRACT

Background: Observational studies raise concern about a potential link between injectable progestin contraceptive use and HIV acquisition risk. This possible link is particularly relevant in sub-Saharan Africa where HIV risk is high and the method mix is skewed toward injectables. We developed the Planning for Outcomes (P4O) model (https://planning4outcomes.ctiexchange.org/) to predict changes in maternal and child health (MCH) and HIV outcomes that could occur if the proportion of injectables in the method mix is changed.

Methods: P4O incorporates evidence-based assumptions to predict yearly changes in unintended pregnancies, morbidity/mortality, HIV infections (women and infants), and anticipated health care costs associated with changing the proportions of injectable users in 22 selected countries. Users of this model designate all countries or a subset and adjust inputs including percentage of injectable users who discontinue, percentage of discontinuers who begin use of an alternative method, hazard ratio for HIV infection with injectable use, method mix used by injectable discontinuers, annual probabilities of method-specific pregnancy and mother-to-child transmission of HIV, condom effectiveness against HIV, risk of HIV during pregnancy, and HIV incidence among women of reproductive age.

Results: Illustrative results from all sub-Saharan African countries combined and from selected countries demonstrate the potential of P4O to inform program planning and procurement decisions. In countries with high use of long-acting reversible contraception, the removal of injectables from the method mix is associated with improvement in MCH and HIV indicators if most injectable users switch to more effective methods (e.g., implants). In countries with high use of short-acting methods (e.g., condoms), the model predicts mostly negative MCH outcomes.

Conclusions: Policy makers and program planners may use P4O to inform programming and policy decisions. In all scenarios, programmatic preparation to accommodate changes to the contraceptive method mix, considerations of how the individual desires of women will be addressed, and potential burden of anticipated MCH-related costs warrant advanced consideration.

BACKGROUND

Access to a range of contraceptive methods is essential for voluntary family planning programs worldwide. Recent analyses demonstrate that for each additional method accessible to at least half the population in a given country, contraceptive use may increase by as much as 8%. In many countries, however, the contraceptive method mix is skewed toward a few methods, with the progestin-only injectable contraceptive depot medroxyprogesterone acetate (DMPA) dominating the method mix in many sub-Saharan African countries.

Observational studies raised concern for a potential link between the use of progestin-only injectable contraceptives, particularly DMPA, and the risk of HIV acquisition. A recent systematic review estimated DMPA users may have a 40% increased risk of HIV acquisition compared with nonusers of hormonal contraception. The authors considered these observational studies to be “...informative but with important limitations to acknowledge that all studies to date are vulnerable to residual or uncontrolled confounding.”

In response to these concerns, the World Health Organization (WHO) recently changed the Medical
Eligibility Criteria for Contraceptive Use (MEC) for progestin-only injectable use among women at high risk for HIV from a 1 (no restrictions) to a 2 (advantages generally outweigh the risks). With results expected in mid-2019, the randomized clinical trial, the Evidence for Contraceptive Options and HIV Outcomes (ECHO) Study, will provide high-quality data on the relative risks of HIV acquisition among African women randomized to use DMPA, the progestin implant Jadelle, or a copper intrauterine device (IUD).

Any potential link between DMPA use and HIV acquisition may critically affect many developing countries, particularly in sub-Saharan Africa where HIV risk is high and the contraceptive method mix is heavily skewed toward injectable progestins. Policy makers in countries with a high HIV burden may inappropriately choose to restrict injectable availability and provision if a significant association is confirmed. While restricting injectable contraceptive availability may lead to fewer HIV infections, this benefit may be offset if women who stop using injectable contraceptives have an unintended pregnancy and are at risk of negative sequelae, including maternal mortality. Moreover, if women with HIV infection who were using injectables move to less effective methods, the number of children born with HIV or acquire HIV in infancy could increase. Preparing for the impact of these theoretical changes in key maternal and child health (MCH) and HIV outcomes, as well as the programmatic desire to ensure women who discontinue injectables are provided alternative method choices to meet their needs, is integral to effective program planning and product procurement.

The literature contains several predictive models that examine the impact of a positive association between injectable hormonal contraceptive use and HIV acquisition risk on MCH and HIV outcomes in sub-Saharan Africa. Jain used data from sub-Saharan African countries on competing risks of unwanted birth and HIV acquisition associated with the use of various contraceptive methods to model ratios of additional unwanted births and additional maternal deaths per 100 HIV infections averted. Similarly, Butler et al. explored country-level effects of reducing injectable hormonal contraceptive use among women of reproductive age on the number of HIV infections, live births, and resulting net consequences on HIV/AIDS deaths and maternal mortality. Lastly, Rodriguez et al. developed a decision-analytic model to compare the benefits and risks of progestin-only injectable use on competing risks of maternal mortality and HIV acquisition on life expectancy in 9 African countries. Our model builds upon these existing tools to offer users an interactive, freely available online interface with adjustable inputs to predict a wide variety of MCH, HIV, and health cost-related outcomes (Box). This tool, Planning for Outcomes (P4O), is available at https://planning4outcomes.citiesexchange.org/.

We developed P4O to enable model users to estimate the impact of changing the amount of injectable progestin use (as a proportion of the method mix) on key MCH and HIV outcomes. P4O is an interactive tool that facilitates policy and program planning decisions and enables countries to better prepare for theoretical anticipated changes. Although P4O will show what is expected to happen mathematically if changes to the method mix occur, programmatic decisions about method changes should be driven by a desire to ensure women who access these programs can make voluntary and informed individual choice. Estimates derived from P4O help highlight the programmatic challenges that may arise when a preferred method of contraception is removed from the contraceptive method mix, including, but not limited to, the potential need for additional training of providers. Addressing these challenges proactively is essential to continuing to provide family planning services that are truly guided by voluntarism and informed choice.

## METHODS

### Model Overview

P4O is an interactive deterministic model that predicts yearly changes in key MCH and HIV indicators for all women of reproductive age (ages 15–49) based on an assumption about the hazard ratio (HR) for HIV acquisition among injectable progestin users, changes to the proportion of injectable progestins in the contraceptive method mix, and redistribution of users to remaining country-specific or a user-specified method mix. In total, 22 countries are included in this model (Figure 1). We included 15 countries with high injectable progestin use as a proportion of the modern contraceptive method mix (≥25%) and an adult HIV prevalence greater than 1%, and an additional 7 countries with either high HIV prevalence or high injectable progestin use. The user can run the model to examine results by individual country, by all countries, or by all sub-Saharan African countries. Users can also adjust the presumed HR for HIV acquisition among injectable progestin users. To date, the literature does not suggest an...
### BOX. What Is Unique About Planning for Outcomes?

Planning for Outcomes (P4O) is an interactive, freely available online tool that allows users to adjust inputs to model key outputs pertaining to maternal and child health and HIV. It is distinct from previous modeling exercises in several notable ways:

- **P4O** incorporates method-specific pregnancy rates and country-specific method mixes.
- **P4O** includes all women of reproductive age, not just those who are married or in union.
- **P4O**’s web-based interface allows users to adjust a variety of inputs, including:
  - Country or region (i.e., all countries modeled, only sub-Saharan African countries, or individual countries)
  - Assumed hazard ratio for HIV infection among injectable users relative to no contraceptive method
  - Proportion of injectable users who adopt other methods
  - How women are reallocated to the existing method mix
  - Additional inputs (i.e., HIV incidence, maternal-to-child transmission, method effectiveness, risk of HIV during pregnancy)
- **Models a wide variety of outcomes**, including (but not limited to):
  - Unintended pregnancies
  - Live births
  - Induced abortions
  - Unsafe abortions (subset of induced abortions)
  - Maternal deaths
  - HIV infections (among women of reproductive age)
  - Children with HIV (from maternal to child transmission)
  - Maternal and neonatal health costs

We use the term “reallocate” to refer to the mathematical reallocation of individuals. Women make individual voluntary and informed decisions.

---

association with HIV acquisition among users of other hormonal methods of contraception or the nonhormonal copper IUD. As such, we only allow for modifications to the risk of HIV among injectable contraceptive users.

**Flow (Inputs/Outputs)**

After designating a country context, users may modify key inputs including *(Figure 2)*:

1. The HR for HIV acquisition with injectable progestin use (compared with no method)
2. Percentage of injectable progestin users who stop the method
3. Percentage of users reallocated to the remaining country-specific method mix (When we use the term “reallocate” to describe movement of injectable progestin users to other methods, we are referring to the mathematical reallocation of individuals to compute model results. We expect women will make individual voluntary and informed decisions on their contraception use.)
4. Inclusion or exclusion of sterilization

To generate country-specific method mix values, we used the most up-to-date national survey data available from a variety of sources (see Supplement). We established an HR range for HIV acquisition with injectable progestins of 0.5 to 5.0 to ensure implausible results would be avoided. We reasoned that injectable progestin users would reallocate to 1 of 3 options when they discontinue the method. In option 1, the model redistributes injectable progestin discontinuers in proportion to the existing, country-specific distribution of other modern methods, after excluding injectables. In option 2, the model redistributes discontinuers in proportion to the existing distribution of other modern methods, after excluding injectable progestins and sterilization (i.e., permanent methods such as bilateral tubal ligation and vasectomy). In option 3, the user may specify the method mix of non-injectable methods to which discontinuers are redistributed.

Key MCH and HIV-related outputs, including yearly changes in unintended pregnancy, morbidity, and mortality, are displayed in graphical or tabular form. A few outputs such as added maternal and neonatal health costs and the percentage of reallocated previous injectable users needed to balance pregnancy outcomes based on the defined method mix (“break-even point”) are summarized at the bottom of the impact panel.
Yearly changes in unintended pregnancy encompasses live births, abortions, and unsafe abortions (calculated as a proportion of all abortions using regional data). Yearly changes in morbidity and mortality comprise maternal deaths, HIV infections among women of reproductive age, and children with HIV due to maternal-to-child transmission (MTCT).

Guidance on Progestin-Only Injectable Contraception

Currently, the WHO MEC aggregates guidance for women at high risk for HIV for the 3 progestin-only injectables (intramuscular [IM] DMPA, subcutaneous [SC] DMPA, and norethindrone/norethisterone enanthate [NET-EN]). A recent review of available data on DMPA and NET-EN suggests that these 2 methods should be disaggregated in the WHO MEC guidance because they have important differences that “may plausibly result in differential impact on HIV susceptibility in women.” No data exist to support such a recommendation on the disaggregation of DMPA-IM and DMPA-SC. While we note the recommendations in this recent review, to be consistent with the current WHO MEC guidance, P4O does not distinguish between the different types of progestin-only injectables (henceforth “injectables”).

Key Assumptions

P4O models outcomes based on a series of evidence-based assumptions. For our MCH indicators, we assumed all women have the annual probability of pregnancy associated with the method to which they are redistributed, and contraceptive prevalence data are consistent with the most recent national survey. We used estimates derived from Marie Stopes International’s Impact 2 Calculator and “Adding It Up” publication to quantify impact on MCH indicators including maternal mortality, MCH health costs, and probabilities of live births and abortions. Furthermore, we assumed women using modern contraception or those with an unmet need due to withdrawal of injectables have at most 1 unintended pregnancy per year and stop using contraception while pregnant. In the event of unintended pregnancy, we assumed women contributed either 12 months of risk-time during pregnancy and postpartum if pregnancy resulted in a live birth, or 6 months of risk-time if the pregnancy did not result in live birth.

For our HIV indicators, most default assumptions (except country-specific HIV prevalence) are modifiable by the user. We assumed HIV prevalence, antiretroviral treatment (ART), and MTCT data are consistent with point estimates provided by AIDSInfo. For each country, we set the pooled HIV incidence among women using contraception to a fixed fraction (default value: 10%).
of the HIV prevalence among women of reproductive age. Due to lack of HIV incidence data among women of reproductive age who use modern contraception, we chose to default to the assumption used by Butler et al.\textsuperscript{7} that incidence is 10% of prevalence in a stable epidemic. We assumed condom users have additional protection against HIV (default value: 85% effective).\textsuperscript{6} Lastly, we assumed no differential risk of HIV during pregnancy, except due to discontinuation of condoms or injectable use. Some existing data suggest an increased risk of HIV acquisition during pregnancy; thus, this value is modifiable.\textsuperscript{20}

**Uncertainty of Estimates**

The model estimates are based on a large array of input parameters, all of which are associated with varying degrees of uncertainty. We recommend that the user explore this uncertainty by varying the model assumptions such as the HR for HIV associated with injectable use, the proportion of injectable users who would stop using the method, and the proportion who would adopt a new method. The assumption about HIV incidence among women of reproductive age using modern contraception and MTCT probabilities are also important drivers of model results, and these can be modified by the user, as well.

**Underlying Model Structure**

In order to concisely explain the model, here we describe only the underlying model structure for determining MCH and HIV-related outcomes when injectable discontinuers are reallocated to methods in proportion to the existing, country-specific distribution of other modern methods, after excluding injectables. As indicated earlier,
however, the model allows an option of defining the method mix alternatively.

### Determining Estimates for Contraceptive and MCH Measures

The model assumes there are \( N_{\text{WRA}} \) women of reproductive age, and that \( P_{\text{INJ}} \) is the proportion using injectable contraception. We wish to estimate the change in the number of pregnancies per year if a proportion \( Y \) of the \( N_{\text{WRA}} \cdot P_{\text{INJ}} \) injectable users stopped using the method, when we further assume a proportion \( X \) of those who stop take up a replacement modern method. Among previous injectable users shifting to a new method, the proportion switching into each type is:

\[
Q_i = P_i / (P_{FS} + P_{MS} + P_{OC} + P_{IUD} + P_{MC} + P_{VB} + P_{IMP} + P_{OTH}),
\]

\[
i \in \{FS, MS, OC, IUD, MC, VB, IMP, OTH\},
\]

where the terms in the denominator denote the current proportion (P) using female sterilization (FS), male sterilization (MS), oral contraceptive pills (OC), IUD, male condoms (MC), vaginal barriers (VB), implants (IMP), or other modern methods (OTH).

Next, the model allows \( PP_j \), where \( j \) is \( \{NM, FS, MS, INJ, OC, IUD, MC, VB, IMP, OTH\} \), denoting the yearly probability of pregnancy when using no method (NM) and so forth. Based on this assumption, the change in the expected number of pregnancies per year, if a proportion \( Y \) of injectable users stop using the method and a proportion \( X \) of those who stop adopt a new method, is given by:

\[
NP_{\text{diff}} = N_{\text{WRA}}P_{\text{INJ}} Y (1 - X) PP_{NM} + N_{\text{WRA}}P_{\text{INJ}} Y X (PP_{FS}Q_{FS} + PP_{MS}Q_{MS} + PP_{OC}Q_{OC} + PP_{IUD}Q_{IUD} + PP_{MC}Q_{MC} + PP_{VB}Q_{VB} + PP_{IMP}Q_{IMP} + PP_{OTH}Q_{OTH}) - N_{\text{WRA}}P_{\text{INJ}} Y PP_{INJ},
\]

where the first line is the expected number of pregnancies among women switching to no method, the subsequent added variables (from the second line through the fourth line) amount to the expected number of pregnancies among women switching to the existing method mix, and the last line is the expected number of pregnancies that would have occurred among injectable users had they not stopped using the method. All other pregnancy-related indicators are obtained by multiplying \( NP_{\text{diff}} \) by the appropriate factor (i.e., the chance an unintended pregnancy leads to a live birth, abortion, unsafe abortion, maternal death, or additional maternal and neonatal health care costs/year).

### Determining Estimates for HIV-Related Measures

To determine outcomes for HIV-related measures in scenarios in which the HR for HIV acquisition with injectable use is greater or less than 1.0, the user must first input an assumption about the overall incidence of HIV among nonpregnant women using modern contraception (denoted \( I_{\text{HIV}} \)). Once \( I_{\text{HIV}} \) is specified, \( P_{\text{4O}} \) calculates distinct HIV incidence values for condom users, injectable users, and users of methods besides condoms or injectables. The model denotes the HR for injectable use versus any method besides condoms as \( HR_{\text{INJ}} \), and the HR for condoms versus any method besides injectables as \( HR_{\text{MC}} \). Then, the incidence of HIV among users of any method besides condoms or injectables is approximated as:
where \( Q_{\text{INJ}} \) is the proportion of the existing method mix, which is injectables, and \( Q_{\text{MC}} \) is the proportion of the method mix, which is male condoms. Then the incidence of HIV among injectable users is:

\[
I_{\text{HIV}}^{\text{INJ}} = \frac{I_{\text{HIV}}}{\{(1 - Q_{\text{INJ}} - Q_{\text{MC}}) + HR_{\text{MC}}Q_{\text{MC}} + HR_{\text{INJ}}Q_{\text{INJ}}\}}.
\]

and the incidence of HIV among condom users is:

\[
I_{\text{HIV}}^{\text{MC}} = \frac{I_{\text{HIV}}}{HR_{\text{MC}}}.
\]

To compute the yearly change in the number of women becoming infected with HIV, if a proportion \( Y \) of current injectable users are withdrawn from the method and a proportion \( X \) of those who stop adopt a new method, we also need to know the prevalence of HIV (PREV\(_{\text{HIV}}\)) (since only those not already infected can become newly infected); the HR for HIV when a woman is pregnant or in the first few months postpartum (HR\(_{\text{PREG}}\)) (since we want to allow for the possibility that HIV acquisition risk changes during this period); and the chance that an unintended pregnancy results in a live birth (\( F_T \)) (since how long a pregnant woman is at differential risk of HIV will depend on whether she carries to term and has a live birth). P\(_4\)O makes the simplifying assumption that women who become pregnant and carry the pregnancy to term contribute up to 1 year of HIV risk while pregnant; women who become pregnant but do not carry to term contribute 6 months of risk while pregnant and 6 months while not pregnant; women stop using their method (including condoms) while pregnant; and women who do not become pregnant contribute 1 year of HIV risk using their contraceptive method. P\(_4\)O then computes the expected change in the yearly number of women becoming infected with HIV as:

\[
N_{\text{WRA}P_{\text{INJ}}}Y(1 - \text{PREV}_{\text{HIV}}) \times \left\{ (1 - X)((1 - PP_{\text{INJ}})(1 - \exp(-I_{\text{HIV}}^{\text{INJ}})) + PP_{\text{INJ}}(F_T(1 - \exp(-I_{\text{HIV}}^{\text{INJ}})HR_{\text{PREG}})) + (1 - F_T)(1 - \exp(-0.5 \cdot I_{\text{HIV}}^{\text{INJ}}(1 + HR_{\text{PREG}})))) \\
+ X \sum Q_i((1 - PP_i)(1 - \exp(-I_{\text{HIV}}^{\text{INJ}})) + PP_i(F_T(1 - \exp(-I_{\text{HIV}}^{\text{INJ}})HR_{\text{PREG}})) + (1 - F_T)(1 - \exp(-0.5 \cdot I_{\text{HIV}}^{\text{INJ}}(1 + HR_{\text{PREG}})))) \\
+ XQ_{\text{MC}}((1 - PP_{\text{MC}})(1 - \exp(-I_{\text{HIV}}^{\text{MC}})HR_{\text{MC}})) + PP_{\text{MC}}(F_T(1 - \exp(-I_{\text{HIV}}^{\text{MC}})HR_{\text{PREG}})) + (1 - F_T)(1 - \exp(-0.5 \cdot I_{\text{HIV}}^{\text{MC}}(HR_{\text{MC}} + HR_{\text{PREG}})))) \\
- ((1 - PP_{\text{INJ}})(1 - \exp(-I_{\text{HIV}}^{\text{INJ}})HR_{\text{INJ}})) + PP_{\text{INJ}}(F_T(1 - \exp(-I_{\text{HIV}}^{\text{INJ}})HR_{\text{PREG}})) + (1 - F_T)(1 - \exp(-0.5 \cdot I_{\text{HIV}}^{\text{INJ}}(HR_{\text{INJ}} + HR_{\text{PREG}})))) \right\}.
\]

where the sum in the third row indexes women adopting \{FS, MS, OC, IUD, VB, IML, OTH\} (with or without becoming pregnant), and the last row captures the number of new infections that would have occurred among the \( N_{\text{WRA}P_{\text{INJ}}}Y(1 - \text{PREV}_{\text{HIV}})\) women who had not stopped using injectables.

To estimate the number of additional children born with HIV if injectable use is reduced, the model determines what percentage of the extra live births were to women with HIV, what percentage of women acquire HIV while pregnant, and the probability infection is transmitted to the child. For the latter, the model must consider the percentage of women with HIV who are on ART (\( \text{PREV}_{\text{ART}} \)), the risk of transmitting HIV to a child when on daily ART (\( P_{\text{ART}}^T(\text{ART}) \)), and the risk of transmitting HIV to the child when not on ART (\( P_{\text{ART}}(\text{ART}) \)). Then, the excess number of children born with HIV is given by:

\[
NP_{\text{diff}} \cdot P_{\text{LB}} \cdot \left\{ \text{PREV}_{\text{HIV}}[\text{PREV}_{\text{ART}} \cdot P_{\text{ART}}^T(\text{ART}) + (1 - \text{PREV}_{\text{ART}}) \cdot P_{\text{ART}}(\text{ART})] + (1 - \text{PREV}_{\text{HIV}}) \cdot \exp(-I_{\text{HIV}}^{\text{INJ}})HR_{\text{PREG}})[\text{PREV}_{\text{ART}} \cdot P_{\text{ART}}^T(\text{ART}) + (1 - \text{PREV}_{\text{ART}}) \cdot P_{\text{ART}}(\text{ART})] \right\}.
\]

### RESULTS

#### Illustrative Scenarios

Since P\(_4\)O is interactive and allows the user to adjust multiple inputs simultaneously, we have selected 4 illustrative examples to demonstrate how the model operates and its potential to inform programmatic decisions. We selected 3 countries with distinct contraceptive method mixes and HIV scenarios—Ethiopia, South Africa, and

Global Health: Science and Practice 2019 | Volume 7 | Number 2
Zimbabwe—to demonstrate the impact of changing the proportion of injectables in settings with varied short-acting and long-acting method use. Additionally, we modeled aggregated outcomes for all 20 of the included countries in sub-Saharan Africa because we predicted that a positive association between injectable use and HIV acquisition would critically impact MCH and HIV indicators in these countries.

In all 4 scenarios, we assumed an HR for HIV with injectable use of 1.4, a 75% discontinuation of injectables, and 25% reallocation proportional to the current country-specific modern method mix after excluding permanent methods (sterilization). All other modifiable inputs remained at the default values. We chose an HR of 1.4 based on the current literature and included the other parameters to model scenarios with demonstrable impact on MCH and HIV indicators. We assumed that if restrictions were placed on injectable contraceptive use, 75% of users would stop the method, and we assumed fewer than 50% would select a new method due to likely real-life programmatic challenges in responding to rapid increases in the resulting method demand. A change in the contraceptive method mix due to women choosing to move from short- to long-acting irreversible contraceptives (LARCs) may present commodity and provider-related challenges (i.e., skill set and availability of providers to provide method). Lastly, we excluded permanent methods because bilateral tubal ligation and vasectomy procedures account for a minority of the method mix in most countries modeled. Model outputs are summarized for each example below and in Table 2.

**Ethiopia**
Ethiopia exemplifies a country setting with relatively low HIV prevalence (1.3% among women of reproductive age) and high use of LARCs (i.e., implants and IUDs). Currently, injectables make

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Ethiopia</th>
<th>South Africa</th>
<th>Zimbabwe</th>
<th>All sub-Saharan Africa</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Yearly change in unintended pregnancies, No.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregnancies</td>
<td>824,933</td>
<td>744,963</td>
<td>51,341</td>
<td>3,565,329</td>
</tr>
<tr>
<td>Live births</td>
<td>464,583</td>
<td>374,107</td>
<td>28,914</td>
<td>1,911,731</td>
</tr>
<tr>
<td>Abortions</td>
<td>243,121</td>
<td>269,123</td>
<td>15,131</td>
<td>1,155,684</td>
</tr>
<tr>
<td>Unsafe abortions</td>
<td>184,976</td>
<td>71,163</td>
<td>11,512</td>
<td>749,343</td>
</tr>
<tr>
<td><strong>Yearly change in morbidity and mortality, No.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women with HIV</td>
<td>-1,227</td>
<td>-22,866</td>
<td>-994</td>
<td>-44,450</td>
</tr>
<tr>
<td>Infants with HIV</td>
<td>988</td>
<td>5,560</td>
<td>329</td>
<td>16,068</td>
</tr>
<tr>
<td>Maternal deaths</td>
<td>2,750</td>
<td>883</td>
<td>167</td>
<td>12,062</td>
</tr>
<tr>
<td><strong>Reallocated DMPA users per method, No.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pill</td>
<td>149,344</td>
<td>152,698</td>
<td>30,606</td>
<td>742,767</td>
</tr>
<tr>
<td>IUD</td>
<td>43,559</td>
<td>27,486</td>
<td>453</td>
<td>179,190</td>
</tr>
<tr>
<td>Male condom</td>
<td>26,965</td>
<td>357,313</td>
<td>4,761</td>
<td>823,462</td>
</tr>
<tr>
<td>Implant</td>
<td>495,739</td>
<td>97,727</td>
<td>9,182</td>
<td>1,256,796</td>
</tr>
<tr>
<td>Other</td>
<td>37,336</td>
<td>6,108</td>
<td>340</td>
<td>160,046</td>
</tr>
<tr>
<td><strong>Additional maternal and neonatal health care costs per year, US$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$35,112,766</td>
<td>$118,779,370</td>
<td>$2,185,284</td>
<td>$249,429,105</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** We are updating the P4O model as new data become available. Thus, results produced may be different from results displayed in this table.

**Abbreviations:** DMPA, depot medroxyprogesterone acetate; HR, hazard ratio; IUD, intrauterine device; MCH, maternal child health.

**Assumptions:** HR for HIV with DMPA=1.4; 75% of injectable users discontinue; 25% reallocate to other methods according to country-specific method mix after excluding permanent methods (sterilization); other parameters set to default.

**Included sub-Saharan African countries:** Botswana, Côte d’Ivoire, Ghana, Mozambique, Nigeria, South Sudan, Burundi, Ethiopia, Kenya, Lesotho, Liberia, Malawi, Namibia, Rwanda, South Africa, eSwatini, Tanzania, Uganda, Zambia, and Zimbabwe.

**Other methods include emergency contraception, Lactational Amenorrhea Method, Standard Days Method, and other modern methods.**
up 63.0% of the method mix; LARCs make up 26.1%; and short-acting methods other than injectables (i.e., contraceptive pills or male condoms) make up 10.3% of the method mix. In our modeled scenario, nearly 500,000 previous injectable users switch to implants and nearly 150,000 switch to contraceptive pills. In this scenario, over 75% of discontinuers shift to a more effective LARC within the existing method mix. In the setting of low HIV prevalence, the impact of these changes in the method mix on HIV acquisition among women is lower than in countries with a high HIV prevalence, with approximately 1,200 fewer women acquiring HIV. Additionally, the model predicts a yearly increase of approximately 243,000 abortions, nearly 76% of which are anticipated to be unsafe.

South Africa

In contrast to Ethiopia, South Africa is notable for its high HIV prevalence—23.8% among women of reproductive age—and mixed use of long- and short-acting contraceptive methods. Injectable contraceptives account for 47.3% of the modern contraceptive method mix, followed by male condoms (24.5%), contraceptive pills (10.5%), and sterilization (8.8%). Among method users, 8.6% use a LARC for contraception. In this modeled scenario, most injectable discontinuers shift to short-acting methods; the model predicts approximately 357,000 and 153,000 discontinuers will switch to male condoms and contraceptive pills, respectively. The model predicts nearly 750,000 unintended pregnancies, over 370,000 live births, and over 269,000 abortions. Further, in the setting of high HIV prevalence, the changes to the method mix are expected to result in approximately 23,000 fewer cases of HIV acquisition among women. Of note, the additional maternal and neonatal health care costs per year are substantially higher than the other 2 countries modeled. Cost estimates, derived from both direct costs, such as personnel time, commodities, medical care, and counseling, and indirect costs, such as program management, health education, advocacy, and infrastructure improvements, are significantly higher in the southern African region.15

Zimbabwe

The prevalence of contraceptive pill use in Zimbabwe creates a unique scenario for modeling outcomes. Contraceptive pills currently account for 56.5% of the modern contraceptive mix, followed by implants (16.9%), injectables (15.1%), and male condoms (8.8%). Among contraceptive users, 17.7% rely on LARCs for contraception. The HIV prevalence among women of reproductive age is relatively high, 16.1%, although substantially lower than that of South Africa. Current injectable use is also the lowest among the 3 countries modeled. In this scenario, most women discontinuing injectable use switch to using a less effective contraceptive method (pills)—more than 3 times as many as those predicted to switch to implants. As women move to less effective methods, unintended pregnancy is expected to increase significantly, with over 51,000 additional unintended pregnancies expected. However, in Zimbabwe, where HIV prevalence is high but injectable use is modest, adjusting the contraceptive use in this scenario results in nearly 1,000 fewer female cases of HIV acquisition.

All sub-Saharan African Countries

In sub-Saharan Africa, there are an estimated 153,113,000 women of reproductive age, with an overall HIV prevalence of 7.1%. Injectables dominate the method mix in aggregated sub-Saharan African countries, with 41.8% of contraceptive users using injectables, followed by implants (16.4%), male condoms (16.1%), pills (14.5%), sterilization (5.1%), and IUDs (2.8%). In this scenario, using 20 countries in the region, the model predicts most reallocated injectable discontinuers will switch to implants (approximately 1,257,000 women) or to male condoms (approximately 823,000 women). In this scenario, approximately 44,000 fewer HIV infections among women are expected; however, this impact is underscored by a predicted 3,565,000 additional pregnancies and nearly 2,000,000 abortions. These outcomes in turn affect the overall added maternal and child health care costs, predicted to total at approximately US$249,429,000.

Discussion

This P4O model serves as a planning tool for policy makers and program planners to input realistic country-specific scenarios and use results to guide contraceptive programming and policy-related decisions. Additionally, the online interface of the model, along with the addition of a variety of instructional materials, makes P4O approachable, accessible, and easily adjustable by a diverse range of users. In most plausible injectable redistribution scenarios, any predicted population-level benefits of reduced HIV incidence that occur by remov-
Injecting injectables from the contraceptive method mix would need to be balanced against the negative public health impacts expected for other outcomes, including increases in unintended pregnancies, abortions, maternal deaths, and HIV infections in children. These findings reflect those in other published modeling work on MCH and HIV outcomes related to injectable contraceptive discontinuation. Rodríguez et al. stated:

...removal of (progestin-only injectable) contraception from the market without effective and acceptable contraception replacement would have a net negative effect on maternal health, life expectancy, and mortality under a variety of scenarios.

However, the outcome of redistribution scenarios is highly dependent on the contraceptive method mix, with more encouraging outcomes expected when women have access to LARCs in the method mix. Increasing the availability of and access to LARC methods may mitigate the public health impact of restricting injectables if DMPA is found to be significantly associated with HIV acquisition. However, changing the contraceptive method mix while ensuring informed and voluntary choice is not a simple task and takes advance preparation because providing increased access to LARCs has significant programmatic, financial, and logistical challenges.

In the scenarios modeled, one may highlight the relative impact of reallocation on potential challenges in product procurement. In each country setting, the demand for rapid procurement of methods to account for the number of previous injectable users switching to other methods is substantial and worthy of advanced consideration. In Ethiopia, for instance, the model predicts nearly half a million previous injectable users will shift to implants. Aggregated results from all 20 sub-Saharan African countries reveal similar trends, with more than 1,257,000 previous injectable users predicted to switch to implants. Family planning programs may use this information to determine whether key factors such as local demand and knowledge, supply, service provision, access to removal services, and other implicated costs are adequately addressed on the timeline needed to prepare for this transition.

Apart from the model’s implications in the setting of HIV acquisition risk, P4O may also serve as a tool for understanding the impact of method skew. In many countries, 50% or more contraceptive users rely on a single method for contraception. By this convention, both Ethiopia and Zimbabwe demonstrate contraceptive use patterns consistent with method skew. South Africa comes close with nearly 50% of the method mix attributable to injectable contraceptives. When most users rely on a single method, it may reflect supply-chain-related challenges in which programs only offer 1 or 2 contraceptive methods rather than the full range of those available. Method skew is attributable to many factors, including but not limited to client characteristics (i.e., age or life stage, desire for limiting versus spacing births), method characteristics (i.e., cost, ease of use, popularity), history (i.e., length of time since introduction of method), provider bias, and policies and programs more broadly. Although positive method characteristics may influence skew, heavily relying on a few contraceptive methods may cause a myriad of downstream challenges if there are sudden, mass shifts between methods. These trends are present in our modeled scenario, which reveal most users in each country setting will disproportionately move from injectables to 1 or 2 methods, namely a combination of contraceptive pills, implants, or male condoms.

Countries need to closely examine their family planning programs to ensure they are prepared to cope with these potential shifts while upholding and advancing volunteerism as well as broad and informed method choice for all clients.

Strengths and Limitations of the P4O Model
To our knowledge, P4O is the first tool of its kind to interactively model the impact of changing the proportion of injectable contraceptive users in the contraceptive method mix. This issue is highly pertinent to the current family planning landscape as we await the ECHO study results. Our model uses evidence-based assumptions and rigorous methodology to model outcomes based on best available estimates for maternal and HIV-related indicators. Further, although the model is currently limited to outcomes specifically related to injectable discontinuation and reallocation, it can be modified to include discontinuation and reallocation of other contraceptive methods and updated data. As hazard ratios for HIV acquisition with other contraceptive methods are published in the literature, we plan to incorporate any risks associated with these methods in future iterations. Lastly, the online and interactive nature of this model facilitates greater accessibility and utility among broad audiences.

P4O is intended to help policy makers, family planning and HIV program planners, and other stakeholders understand the potential impact of
a change in current injectable contraceptive prevalence on pregnancy and HIV outcomes. As with any model, however, flawed or implausible assumptions can lead to flawed or implausible outputs. P4O possesses several additional key limitations. First, the model is limited to 22 countries; further shifts in the method mix, or new interventions to prevent or treat HIV. Third, the model does not consider the impact of pre-exposure prophylaxis or condom use, which may differ by contraceptive method. Lastly, the model does not distinguish between injectable progestins, which may have different risks. We plan to update P4O if and/or when data differentiating risks among injectable progestins are available.

**CONCLUSIONS**

As the world awaits results clarifying data from a randomized controlled trial on the relationship between DMPA use and HIV acquisition, programs must begin to consider downstream implications of any negative findings. We hope this model attracts and is useful to a variety of users with diverse interests—health care policy makers, ministry of health officials, family planning and HIV program planners, commodities procurers, advocacy groups, funders, and those with general interest in global family planning. Regardless of how different audiences choose to use this tool, we underscore the importance of individual choices when planning for potential outcomes and urge programs to consider these results and their implications when making programmatic and policy-related decisions.

**Acknowledgments:** The authors wish to thank Maggwa Baker, Marylyn Gaffield, Renee Heffron, James Kiarie, Nithya Mani, Jennifer Mason, Kevin Peine, Tabitha Snipipatana, Petrus Steyn, and Sarah Yeiser for their thoughtful feedback in the development of Planning for Outcomes.

**Funding:** This work is made possible by the generous support of the American people through the United States Agency for International Development (USAID), provided to FHI 360 through Cooperative Agreement AID-OAA-A-15-00045.

**Disclaimer:** The content and views expressed by the authors in this publication do not necessarily reflect the views of USAID, the United States Government, or FHI 360.

**Competing Interests:** None declared.

**REFERENCES**


The Extent to Which Performance-Based Financing Programs’ Operations Manuals Reflect Rights-Based Principles: Implications for Family Planning Services

Marie S. Cole, a Victoria Boydell, b Karen Hardee, c Ben Bellows d

Rights principles should be prioritized and more clearly stated in performance-based financing (PBF) guidance and operational documents. Additional research, including development of validated measurement metrics, is needed to help PBF programs systematically align with rights-based approaches to health care including family planning.

ABSTRACT

Recognition is growing that development programs need to be guided by rights as well as to promote, protect, and fulfill them. Drawing from a content analysis of performance-based financing (PBF) implementation manuals, we quantify the extent to which these manuals use a rights perspective to frame family planning services. PBF is an adaptable service purchasing strategy that aims to improve equity and quality of health service provision. PBF can contribute toward achieving global family planning goals and has institutional support from multiple development partners including the Global Financing Facility in support of Every Woman Every Child. A review of 23 PBF implementation manuals finds that all documents are focused largely on the implementation of quality and accountability mechanisms, but few address issues of accessibility, availability, informed choice, acceptability, and/or nondiscrimination and equity. Notably, operational inclusion of agency, autonomy, empowerment, and/or voluntarism of health care clients is absent. Based on these findings, we argue that current PBF programs incorporate some mention of rights but are not systematically aligned with a rights-based approach. If PBF programs better reflected the importance of client-centered, rights-based programming, program performance could be improved and risk of infringing rights could be reduced. Given the mixed evidence for PBF benefits and the risk of perverse incentives in earlier PBF programs that were not aligned with rights-based approaches, we argue that greater attention to the rights principles of acceptability, accessibility, availability, and quality; accountability; agency and empowerment; equity and nondiscrimination; informed choice and decision making; participation; and privacy and confidentiality would improve health service delivery and health system performance for all stakeholders with clients at the center. Based on this review, we recommend making the rights-based approach explicit in PBF; progressively operationalizing rights, drawing from local experience; validating rights-based metrics to address measurement gaps; and recognizing the economic value of aligning PBF with rights principles. Such recommendations anchor an aspirational rights agenda with a practical PBF strategy on the need and opportunity for validated metrics.

INTRODUCTION

Since the 1994 International Conference on Population and Development, the family planning movement has increasingly focused on person-centered or client-centered approaches to ensure individuals and couples can freely determine whether and when to bear children. However, significant barriers remain, inhibiting millions from realizing their reproductive intentions. An estimated 214 million women who want to delay or refrain from childbearing are not using contraception. As such, increasing access to voluntary family planning services is a global health goal, as demonstrated by the Sustainable Development Goals and Family Planning 2020. Performance-based financing (PBF) is a common approach to achieve global health goals, including family planning, because it aims to link efficient provision of high-quality health services with incentives. This approach is now one of several financial strategies of the Global Financing Facility (GFF) in support of Every Woman Every Child, the new flagship mechanism for reproductive, maternal, newborn, child, and adolescent health in 63 high-burden low- and middle-income countries (Box).

Growing interest also exists for operationalizing rights principles in health and development initiatives to better meet client needs and ensure states fulfill their

---

a PAI, Washington, DC, USA.
b Graduate Institute, Geneva, Switzerland.
c What Works Association, Morristown, NJ, USA.
d Population Council, Washington, DC, USA.
Correspondence to Ben Bellows (bbellows@popcouncil.org).
BOX. Definitions for Performance-Based Financing and Rights-Based Approach in Family Planning

What Is Performance-Based Financing?
Performance-based financing is an instrument through which payments are made to providers, health facilities, or local administrative units (e.g., district health offices) for health services conditional on the performance of predefined and verified quantity indicators, adjusted for measures of quality.\(^{10}\)

What Is a Rights-Based Approach to Family Planning?
A rights-based approach to family planning uses a set of human rights standards and principles to guide program assessment, planning, implementation, monitoring, and evaluation that enable individuals and couples to decide freely and responsibly the number and spacing of their children, to have the information and services to do so, and to be treated equitably and without discrimination.

PBF programs could strengthen person- or client-centered care with performance incentives that respect and protect individual rights.

We undertook an evidence mapping of PBF operational documents to assess alignment between PBF programs and rights principles.

OBJS. Definitions for Performance-Based Financing and Rights-Based Approach in Family Planning

What Is Performance-Based Financing?
Performance-based financing is an instrument through which payments are made to providers, health facilities, or local administrative units (e.g., district health offices) for health services conditional on the performance of predefined and verified quantity indicators, adjusted for measures of quality.\(^{10}\)

What Is a Rights-Based Approach to Family Planning?
A rights-based approach to family planning uses a set of human rights standards and principles to guide program assessment, planning, implementation, monitoring, and evaluation that enable individuals and couples to decide freely and responsibly the number and spacing of their children, to have the information and services to do so, and to be treated equitably and without discrimination.

PBF programs could strengthen person- or client-centered care with performance incentives that respect and protect individual rights.

We undertook an evidence mapping of PBF operational documents to assess alignment between PBF programs and rights principles.

OBS. Definitions for Performance-Based Financing and Rights-Based Approach in Family Planning

What Is Performance-Based Financing?
Performance-based financing is an instrument through which payments are made to providers, health facilities, or local administrative units (e.g., district health offices) for health services conditional on the performance of predefined and verified quantity indicators, adjusted for measures of quality.\(^{10}\)

What Is a Rights-Based Approach to Family Planning?
A rights-based approach to family planning uses a set of human rights standards and principles to guide program assessment, planning, implementation, monitoring, and evaluation that enable individuals and couples to decide freely and responsibly the number and spacing of their children, to have the information and services to do so, and to be treated equitably and without discrimination.

Without a rights-based focus on clients, incentives linked to family planning services may stimulate adverse selection, with providers serving only clients who are likely to meet PBF performance standards.\(^{26}\) This gap in addressing client-centered needs, together with concerns about the risks for perverse incentives in PBF, motivated an exploration of the rights in family planning services in PBF program implementation.\(^{11,28,30,31}\)

A recent literature review on performance-based financing policy draws lessons on transitioning donor-supported schemes to the health system in 4 stages (generation, adoption, institutionalization, and expansion) occurring along 5 dimensions (population coverage, service coverage, health system integration, cross-sectional diffusion, and knowledge expansion).\(^{32,33}\) This transition is defined elsewhere as the process by which a PBF project becomes an integral part of the national health system; as noted elsewhere, the process does not have a predetermined outcome.\(^{34}\) Our review of rights principles contributes to this larger discourse by suggesting another dimension to consider for PBF programs—that is, the degree to which they systematically align with a rights-based approach.

A REVIEW OF RIGHTS IN PBF OPERATIONS MANUALS

To assess the degree of alignment between PBF programs and rights principles, we undertook an evidence mapping of PBF operational documents from November to December 2017. These operational documents were not discoverable through published literature databases. The search focused on the World Bank Group’s results-based financing website, which served as a repository for PBF documentation, followed by a request to experts in the online PBF Community of Practice on Google and Collectivity groups to share additional documentation. We were unable to source all existing operational manuals because they were not available in the public domain. The exploration identified 23 relevant documents (one each from 23 countries) in English and French. The manuals were written by country teams most often with
support from the World Bank, although it was not always possible to determine which donors supported the manual production. Likewise, we could not determine the level of government buy-in for specific PBF programs. Government leadership on health policy reforms, including health purchasing such as PBF, is a prerequisite for sustained success in the drive for universal health coverage.\textsuperscript{35} The debate regarding the degree to which PBF is a function of donor and government leadership is complex and beyond the scope of this review.\textsuperscript{36}

The documents were the most recent publicly available versions of PBF operations or implementation manuals. Concepts, procedures, and performance measures were extracted and mapped to 8 rights principles for family planning sourced from global agreements, including Family Planning 2020 (2014),\textsuperscript{12} the World Health Organization (2014),\textsuperscript{13} and a revision of the Family Planning Quality of Care Framework\textsuperscript{37} (Table 1). The definitions and implications for each rights principle provided in Table 1 guided data extraction: any text within the PBF manuals that reflected one or more of the rights principles was selected for analysis. The evidence was extracted to a spreadsheet that tracked the degree to which each rights-based principle was included in each manual. The categories for data extraction were document type, date, each right principle, overlapping/miscellaneous indicators, and mention of family planning.

Data extraction was not exclusively focused on family planning-specific elements, although we gave attention to family planning implications. Identifying rights principles in the documents involved linking the terms for the rights described in Table 1 with specific elements in the operations manuals. These links were not always specific to family planning; for example, the availability of commodities is inclusive but not necessarily specific to contraceptives.

### PBF IMPLEMENTATION DOCUMENTS INCONSISTENTLY INCLUDE RIGHTS PRINCIPLES

Of the 23 implementation manuals reviewed, 21 manuals mention or include family planning. All documents focused largely on the implementation of quality and accountability mechanisms, but few addressed issues of accessibility, availability, informed choice, acceptability, and/or nondiscrimination and equity. Notably, operational inclusion of agency, autonomy, empowerment, and/or voluntarism of health care clients was absent. Table 2 illustrates which principles were included in each of the manuals.

The manuals reviewed were produced between 2009 and 2015, prior to the GFF.\textsuperscript{38–60} As such, they could be updated and improved with guidance from the GFF and partner countries to include rights principles in future editions.

#### Rights Are Partially Recognized in PBF Guidance

Apart from agency, all the rights principles are included across the reviewed PBF implementation documents, albeit to varying degrees. However, gaps are present in the inclusion of some principles. For example, under principles of equity and nondiscrimination, income and ethnicity were addressed by many programs, but only a few addressed youth, a critically disadvantaged and often neglected population. In addition, minimal emphasis was placed on addressing financial or information barriers to health care. For example, the operational manuals from Burkina Faso and Mozambique noted use of “equity bonus” adjustments made to PBF payments as a means to address issues of geographic access that could affect service delivery.\textsuperscript{40,51} Equity, as Chowdhury and colleagues note, is often cited as a PBF program indicator with its own corresponding payment mechanisms.\textsuperscript{28} In Mozambique, the equity bonus is paid based on the quantity of services provided in particular districts by a health facility, and it must be used for structural repairs or reproductive supply purchases. While the decision to focus on districts is one answer to resolving problems of geographic access that could affect service delivery to health care clients who are potentially more disadvantaged or vulnerable, the equity bonus does not address whether those clients are able to afford or understand available services. Burkina Faso’s program provided ample detail for assessing the equity indicator. Some of the criteria used as a measure for the equity bonus include incidence of poverty in the area of service, population density, and distance between health facility and villages served.\textsuperscript{40} A further example of equity and nondiscrimination provided in PBF operational manuals is the payment for some services provided to the very poor in Burkina Faso and Nigeria.\textsuperscript{52} Nevertheless, concerns around the various monetary costs for health care paid by clients—including but not limited to travel and service costs—were not considered in most of the PBF operational manuals analyzed.

#### Client Agency Receives Limited Attention

A rights-based approach centers on clients making voluntary, informed decisions about their
contraceptive use. The principle of informed choice focuses on an individual’s ability to access and readily understand information about a variety of contraceptive methods and their use. This right to information is a distinct right, antecedent to agency—a right to take action, which is grouped with autonomy, empowerment, and voluntarism for purposes of this review. It implies women have the right to make decisions about having children and to act on those decisions in the health care system through knowledge of their right to family planning, voluntarily and free of discrimination, coercion, or violence. The provision of family planning options and complete information about the alternatives is one dimension of agency alongside women’s capacity to articulate and drive their own decisions and to influence others to support them in achieving their goals. Aspects of agency beyond informed choice were completely absent from all the PBF manuals. Other rights principles related to agency, such as informed choice, had limited mentions in the manuals with the exception of those from Senegal and Tanzania.

### TABLE 1. Principles to Guide a Rights-Based Approach to Family Planning

<table>
<thead>
<tr>
<th>Rights Element</th>
<th>Family Planning Program Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accessibility</td>
<td>Geographic, physical, financial, and policy access (i.e., absence of nonmedical eligibility criteria); information is available in the languages and terms people can understand; continuous contraceptive security; suitable hours of operation; service integration to increase access</td>
</tr>
<tr>
<td>Acceptability</td>
<td>Culturally appropriate facilities, methods, and services; community/family support for women’s ability to choose, switch, or stop method of contraception; tolerance of side effects; privacy and confidentiality respected; client satisfaction with services</td>
</tr>
<tr>
<td>Accountability, participation, transparency</td>
<td>Mechanisms exist for community members and family planning clients to provide input and feedback about services, and for health system to investigate and remedy allegations of or confirmed violations of rights; members of the community are involved in planning and monitoring family planning services; good governance and effective implementation, providing an environment that facilitates the discharge of all responsibilities; and the ability to readily access meaningful information, including de-identified data.</td>
</tr>
<tr>
<td>Agency, autonomy, empowerment, voluntarism</td>
<td>Knowledge that one has the right to make decisions about health care; ability to make one’s own decisions independent of system, husband, family, or community pressures; informed, voluntary decision making supported; meaningful participation of clients in program design and monitoring; client-controlled methods offered; supportive community gender norms; women, men, and young people know they can ask for services based on their needs, within their rights</td>
</tr>
<tr>
<td>Availability</td>
<td>Broad choice of methods offered; sufficient and needs-based distribution at functioning service delivery points</td>
</tr>
<tr>
<td>Informed choice, informed decision making</td>
<td>Women and youth and all clients make own decisions about whether and what method of family planning to use, without pressure from anyone, with free access to accurate information they can understand and a range of options to choose from</td>
</tr>
<tr>
<td>Nondiscrimination, equity</td>
<td>Everyone, no matter what group they come from, their age, or any other circumstance, has the same access to quality information and services; everyone is treated fairly and equitably</td>
</tr>
<tr>
<td>Quality (including privacy and confidentiality)</td>
<td>Service providers are well trained and provide safe services, treat clients with respect, provide good counseling, and protect client privacy and confidentiality (ensuring client information cannot be observed by anyone else without client’s consent; ensuring client records are not disclosed); stock a regular supply of contraceptives and all necessary equipment to provide the services clients want</td>
</tr>
</tbody>
</table>

*These rights principles for family planning flow from global treaties, covenants, and conventions that define rights broadly. Sources: Modified from Family Planning 2020 (2014),12 WHO (2014),13 and Kumar et al. (2017).86 Note that the definition of quality also incorporates components from the updated Bruce/Jain Quality of Care Framework for Family Planning.37*
which could result in subtle pressure or coercion to use family planning or to take a certain method. For example, a payment to a facility or provider for each family planning client served could perversely incentivize the provider to recommend short-acting methods, knowing that many clients using these methods will return for refills.\textsuperscript{19,61} Such an incentive could be modified to better align with a rights-based approach. For example, quality checklists and verification processes could be adapted to included validated measures of family planning clients’ opportunity and confidence to ask questions during a consultation, the quality of counseling received, if they received their preferred method, and who was thought to influence their decision making.\textsuperscript{62–65} A recent report described what operationalizing voluntarism could look like; for instance, mechanisms could be developed to monitor for signs of coercive practices, such as setting up a confidential hotline, mobile reporting, or community dialogues.\textsuperscript{61}

### TABLE 2. Rights Principles in 23 Performance-Based Financing Operational Manuals

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tanzania</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Lesotho</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Afghanistan</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Argentina</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Armenia</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Benin</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Burkina Faso</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Burundi</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Cameroon</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Democratic Republic of the Congo</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Djibouti</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Haiti</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Ivory Coast</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Kenya</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Liberia</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Mali</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Mozambique</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Nigeria</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Rwanda</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Senegal</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Sierra Leone</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Tajikistan</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Zambia</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>6</td>
<td>4</td>
<td>23</td>
<td>0</td>
<td>6</td>
<td>4</td>
<td>7</td>
<td>19</td>
<td>21</td>
</tr>
</tbody>
</table>

Abbreviations: FP, family planning; PBF, performance-based financing.
The wide-ranging and complex operationalization of quality remains a challenge for PBF.

Downward Accountability Measures Are Limited
Accountability is the only rights principle addressed in every implementation manual reviewed because it is embedded within the required verification process for PBF programs. However, the accountability measures found in the manuals only relate to internal performance checks to verify service delivery before making payments. These accountability structures connect health care providers to PBF funders and regulators but not to health care clients. Since the manuals focused on supply-side PBF programs and the contracts among funders, regulators, and health service providers, they also concentrated on these actors’ respective roles and responsibilities in relation to the programs’ incentives. The manuals minimally discussed how health care users should be served by the health system. In large part, the health system’s responsibilities to clients, as expressed in the exit interviews, verification observations, and reviews of facility registers, focus on whether the client was served by a facility.

Previous studies of PBF have noted that verification (accountability) procedures, such as confirming poverty status or use of services, may prefer verification over the client’s right to confidentiality and distorts the system’s ability to monitor that client-centered services were delivered. In a telling example, PBF program managers in Mozambique consulted village leaders and community members as they sought to verify an HIV client’s prior visit to clinic. In the process, they infringed the man’s right to confidentiality, imposing significant social harm and costs on him and his household. Greater client and community participation in program design and implementation can help to align PBF accountability procedures with a rights-based approach. There are promising examples from community PBF programs in which individuals have an active role in the business plan, quality of care delivered, and identification of the beneficiaries—increasing the potential of communities providing oversight in health facilities’ activities.

The Concept of Quality Is Incomplete
Quality is a common rights principle used in PBF performance measurement and incentives, as reflected in the PBF documents reviewed. Although the checklists used were structured on the accepted quality of care framework, how quality was emphasized and assessed varied across the manuals and tended to focus on tangible, structural dimensions over process indicators, such as counseling and health outcomes. Quality often overlapped with other rights principles, namely accessibility, availability, and acceptability. The wide-ranging and complex operationalization of quality remains a challenge for PBF, as for many health care programs and health systems. Critical elements, such as the quality of counseling beyond informed choice, received less attention, although they have been shown to have marked effects on increased contraceptive use.

Concentration Is on Supply Over Demand
As is common in PBF programs, the manuals focused on supply-side financial incentives. This focus speaks to PBF design and should not be understood as a unique strategy of the GFF, which encourages countries to consider a wide range of possible approaches from high-level domestic resource mobilization, revenue pooling, and innovative financing mechanisms to improve efficiency and equity in service purchasing. Demand-side incentives were not present in the reviewed implementation manuals. The manuals often frame the PBF theory of change as an empowering, equitable, efficient, and effective approach for service providers, and a strong focus on the client is lacking. Beyond PBF, which largely focuses on supplier incentives, demand-side programs are critical to address barriers faced by clients accessing family planning services.

Limitations
This article is limited to the evidence mapping of PBF operations manuals that were available via web searches and supplemental requests to individuals. The search may have missed other PBF documents and does not include recently developed PBF manuals. Challenges in data extraction included the overlaps in definitions of rights principles. For example, accessibility and availability are necessarily related principles in obtaining family planning services. Identifying the right principle linked to an indicator such as “new contraceptive user” was challenging when the objective underlying each PBF indicator was unclear in the manual.

Another limitation is the exclusion of other kinds of results-based financing programs such as vouchers and conditional cash transfers. As noted by Musgrove, PBF is a health financing instrument that directs conditional payments to providers for health services that meet verified quantity indicators, adjusted for quality. PBF is one of multiple strategies that the GFF and the
Operationalizing Rights Principles in PBF

This review assessed current PBF design and implementation processes, inclusive of contraceptive services, against 8 rights principles. Currently, PBF manuals do not provide practical guidance on how to operationalize rights principles in a systematic way. The emphasis in the reviewed manuals is often on the autonomy of providers and facilities but not that of clients. Moreover, we noted that PBF programs are constrained by what they measure. The lack of metrics to observe certain rights would preclude alignment of PBF to rights-based frameworks. Given the significance of PBF in health care purchasing for many low- and middle-income countries, operationalizing rights principles is an urgent priority. The evolution of PBF for family planning services will have significant implications for the abilities of women and couples to choose if, when, and how many children to bear in their lifetimes, and we must ensure it is rights based, with emphasis on universal self-determination of fertility intentions.

A significant share of PBF technical support is administered by the World Bank, with funding from governments that have signed international agreements intended to ensure the right to the highest attainable standard of health. The GFF Business Plan states that “equity, gender, and rights underpin and are mainstreamed throughout the GFF’s work,” and it provides the foundation for a new concentrated effort to explicitly integrate rights principles into the work of PBF. As the principal in providing operational guidance on the use of rights-based approaches, the World Bank can support nation-states in realizing their commitments to rights principles in PBF initiatives.

Based on this review, we recommend that this guidance include the following:

Recommendation 1: Make the Rights-Based Approach, Centered on the Client, Explicit

Guidance on explicitly integrating a rights-based approach into PBF should start with clarifying how rights are realized within the PBF theory of change, including how such programming supports clients in addition to providers and programs. This explanation would place clients at the heart of services, ensuring they are prioritized, protected, and supported in making their own decisions about contraceptive use, within the design and implementation of PBF programs. Taking a rights-based approach would motivate the inclusion of robust indicators for quality, such as the method information index, and the development of indicators for other rights principles, such as access, choice, voluntarism, and equity. Furthermore, PBF programs built on a rights-based approach would strengthen accountability systems that move beyond financial accountability to also include community-level or social accountability for the services provided. Careful review of PBF indicators is important to ensure that the programs are not infringing on rights, for example, by denying a client informed choice by rewarding providers for provision of certain contraceptive methods over others. This articulation and prioritization of rights principles should be explicit in the guidance throughout the PBF lifecycle, from investment cases to project appraisal documents to PBF operational manuals and the monitoring and verification process, which drives the payment of incentives. The directives should be oriented first and foremost around the client, with attention to expressing quality, informed choice, voluntarism, and other rights principles in clear consistent terms throughout PBF programs. Additionally, relevant technical support should be developed to sensitize and support country teams in developing and implementing their PBF programs.

Recommendation 2: Progressively Operationalize Rights, Drawing From Local Experience

The rights identified in the manuals serve as practical examples to be considered and taken up by other PBF programs. Current and future experiences need to be fully documented, assessed, and disseminated as a core part of PBF guidance. How rights principles are incorporated into implementation will require a thoughtful, iterative approach to document insights while accounting for contextual variation. Determining the design, implementation, and monitoring/verification of a PBF scheme as well as compensation to providers through a rights-based approach will ultimately be mediated through the local context and health system.
Mapping Rights-Based Principles in Performance-Based Financing Programs

**Recommendation 3: Validate Rights-Based Metrics to Address Measurement Gaps**

To ensure a systematic integration of rights into PBF programs, new metrics are necessary to address current gaps in measurement; in particular, metrics for agency and autonomy in the health system are needed. Secondly, the issue of a broad and complex definition of quality is a challenge. For over 25 years, the family planning field has been guided by the Bruce Quality of Care Framework. Recently, Jain proposed a revision of the framework that aligns it with definitions of quality in frameworks for rights-based family planning. The revised family planning quality of care framework could serve as a foundation on which quality in PBF is more clearly centered on rights.

PBF programs would benefit from having a core set of indicators covering the rights principles that they could choose from to measure in their programming. This review and other recent work provide a good start at anchoring rights-based programs and aspects of PBF on metrics. New work is needed to advance a rights-based measurement agenda for family planning services in PBF and propose a list of potential indicators for validation and testing in PBF programs.

**Recommendation 4: Recognize the Economic Value of Aligning PBF With Rights Principles**

The odds of successfully translating the value of a rights-based approach beyond its current community of practice may be improved by appreciating the operational enhancements that a rights-based approach could bring to PBF program performance. In this utilitarian perspective, stakeholders take into account the economic role and value of client voice and agency in PBF programs and the harm to programs that abridge those rights. If PBF services fail to meet clients’ expectations for quality, confidentiality, or agency, they may impose unintentional and unaccounted costs on households. For example, if a PBF program sends field verification teams to a community where they unintentionally disclose the contraceptive use status of an adolescent PBF beneficiary, the stigma experienced by that adolescent represents a potential economic cost to them and may dissuade other young people from seeking care. Providers’ negative attitudes toward adolescents’ contraceptive use are well documented. Similarly, if incentivized services reward facilities for family planning patient volume instead of clients’ informed choice, providers may be incentivized to disregard client choice, which places clients at risk of earlier-than-desired contraceptive discontinuation and imposes personal and social costs. Inattention to equity can result in some groups not receiving services for which they qualify. Although emerging evidence shows that higher client-perceived quality is associated with higher rates of contraceptive continuation, further research is needed to quantify the direct and indirect benefits of respectful, rights-based family planning in the context of performance-based financing.

**Competing Interests:** None declared.

**REFERENCES**

Mapping Rights-Based Principles in Performance-Based Financing Programs


Mapping Rights-Based Principles in Performance-Based Financing Programs


Global Health: Science and Practice 2019 | Volume 7 | Number 2


