GHSP

Dedicated to what works in global health programs

GLOBAL HEALTH: SCIENCE AND PRACTICE

2019 | Volume 7 | Number 4

www.ghspjournal.org

USAID FROM THE AMERICAN PEOPLE

JOHNS HOPKINS Center for Communication Programs

UNIVERSITY OF ALBERTA SCHOOL OF PUBLIC HEALTH

Global Health: Science and Practice

EDITORS

EDITORIAL BOARD

Global Health: Science and Practice (ISSN: 2169-575X) is a no-fee, open-access, peer-reviewed journal published online at www.ghspjournal.org. It is published quarterly by the Johns Hopkins Center for Communication Programs, 111 Market Place, Suite 310, Baltimore, MD 21202. GHSP is made possible by the support of the American People through the United States Agency for International Development (USAID) under the Knowledge SUCCESS (Strengthening Use, Capacity, Collaboration, Exchange, Synthesis, and Sharing) Project. GHSP is editorially independent and does not necessarily represent the views or positions of USAID, the United States Government, or the Johns Hopkins University. For further information, please contact the editors at editorialoffice@ghspjournal.org.

Cover caption: A quadcopter drone delivers medical supplies in Rwanda. © 2018/HealthNews

Table of Contents

December 2019 | Volume 7 | Number 4

EDITORIALS

COMMENTARIES

ORIGINAL ARTICLES

Authors' Response to "Saving Mothers, Giving Life: Don't Neglect the Health Systems Element"

Florina Serbanescu, Claudia Morrissey Conlon, Frank Kaharuza, Masuka Musumali

Glob Health Sci Pract. 2019;7(4):610–611
https://doi.org/10.9745/GHSP-D-19-00313

GLOBAL HEALTH: SCIENCE AND PRACTICE
Dedicated to what works in global health programs
OPEN ACCESS

EDITORIAL

A Tablet-Based Tool for Care During Labor+Attention to System Requirements

Stephen Hodgins[a]

Evidence on using a tablet-based labor decision-support tool suggests the potential for improved practices in labor management. Further rigorous study on these tools is needed to assess the improvements in labor care and outcomes as well as the system requirements needed to achieve such improvements.

See related article by Sanghvi et al.

Over the past decade, it has been heartening to see marked increases in institutional births in many countries. However, this increase has not consistently been associated with corresponding declines in maternal and perinatal deaths.[1–3] In many countries, intrapartum stillbirths and very early newborn deaths remain distressingly common. Safe, supportive, vigilant, and well-coordinated care during labor could significantly reduce the burden of such deaths.

USING THE PAPER PARTOGRAM TO AID LABOR DECISION MAKING

The partogram is a single-page, graphic record used to document key clinical information on the mother and fetus during labor and is intended to be an aid to clinical decision making. Although the World Health Organization no longer recommends using the 4-hour action line (based on 1 cm/hour increase in dilatation through the active phase of labor), it does endorse using the partogram for monitoring the well-being of the mother and newborn and for identifying risks of adverse outcomes.[4] In principle, as designed, partogram use might be expected to contribute to improved outcomes, mediated through the following:

- Improvement of practices that can affect placental perfusion
- More timely identification of complications or risk states and initiation of appropriate action, including transfer or referral
- Better documentation and sharing of information between involved health workers on care given and status of the mother and newborn, at handover or referral

However, the available evidence for such an effect has been disappointing. A Cochrane review by Lavender et al. (2018)[5] found no clear evidence for an association between partogram use and improved practices or outcomes, although they offered more nuanced conclusions in a realist review on the same question.[6] From this review, they concluded that to achieve improved outcomes:

- The partogram and all associated equipment (e.g., manometers for taking blood pressure) need to be reliably available
- Staffing needs to be adequate for patient load
- Clinical management needs to provide ongoing, committed support for partogram use and associated labor care practices, including regular audit and feedback

In short, even clinically sound, well-designed job aids cannot be expected on their own to improve practices and outcomes. Indeed, deficiencies in the use of the original paper-based partogram are well known. In many instances, they are filled in after the fact, information recorded may be inaccurate, and they are often not used for decision making. Even if they are used for decision making, there has tended to be excessive emphasis on rate of cervical dilatation.[6,7]

USING THE TABLET-BASED PARTOGRAPH

The article by Sanghvi et al.[8] in this issue of GHSP reports on a study conducted in Kenya that attempted to address some of these known constraints on the effectiveness of partogram use. This study examined the use of a tablet-based documentation/decision-support tool with the following key features:

- Clinical information can only be entered in real-time, not after the fact.

[a] Editor-in-Chief, Global Health: Science and Practice Journal, and Associate Professor, School of Public Health, University of Alberta, Edmonton, Alberta, Canada.
Correspondence to Stephen Hodgins (shodgins@ghspjournal.org).

- Information entered into the tablet can be monitored by a supervisor, either on- or off-site.
- Automated prompts are given, encouraging supportive practices (e.g., continued ambulation, presence of a labor companion, taking fluids and food).
- Audible alarms/triggers remind the health worker on timing for reassessment and indicate if, based on the algorithms in the app, criteria are met suggesting a risk or complication requiring action. (As the authors note, when guidelines for care in labor are further revised, the algorithms used in this tool can be easily reprogrammed.)

Participants in both intervention and comparison arms were given a 2-day training on care during labor, including partogram use. Those in the intervention arm were given an additional 1-day training on the tablet-based tool, which, in the intervention arm, was subsequently used in place of the paper partogram.

The intervention was well received by participating health workers, and the study provides suggestive evidence for improved practices. However, as the authors acknowledge, there are issues with the design and implementation of the study that preclude definitive claims for impact on health outcomes. First, the study sites in the intervention and comparison arms were not well matched: the intervention arm had 2 tertiary-level hospitals with large patient volumes plus 4 health centers offering basic emergency obstetric and newborn care (BEmONC), and the comparison arm had 4 relatively lower-volume comprehensive emergency obstetric and newborn care hospitals and 2 health centers offering BEmONC. Second, compared with sites in the intervention arm, a significantly lower proportion of health workers in the comparison arm received the 2-day training. Furthermore, data on performance in the comparison arm were more limited because data were drawn only from the completed partograms.

The evidence from this and similar studies is not yet sufficient to recommend large-scale adoption.[9–11] Nevertheless, a strong case can be made for further rigorous study to better characterize what improvements can be achieved in care during labor using such tools and what system requirements must be met to achieve such improvements (such as those identified by Bedwell et al.[6]).

Even clinically sound, well-designed job aids cannot be expected on their own to improve practices and outcomes.

The intervention was well received, and the study provides suggestive evidence for improved practices.

REFERENCES

1. Fink G, Ross R, Hill K. Institutional deliveries weakly associated with improved neonatal survival in developing countries: evidence from 192 Demographic and Health Surveys. *Int J Epidemiol.* 2015;44(6):1879–1888. CrossRef. Medline
2. Kunkel M, Marete I, Cheng ER, et al. Place of delivery and perinatal mortality in Kenya. Semin Perinatol. 2019;43(5):252–259. CrossRef. Medline
3. Gabrysch S, Nesbitt RC, Schoeps A, et al. Does facility birth reduce maternal and perinatal mortality in Brong Ahafo, Ghana? A secondary analysis using data on 119244 pregnancies from two cluster-randomised controlled trials. *Lancet Glob Health.* 2019;7(8):e1074–e1087. CrossRef. Medline
4. World Health Organization (WHO). *WHO Recommendations: Intrapartum Care for a Positive Childbirth Experience.* Geneva, Switzerland: WHO; 2018.
5. Lavender T, Cuthbert A, Smyth RMD. Effect of partograph use on outcomes for women in spontaneous labour at term and their babies. *Cochrane Database Syst Rev.* 2018;8:CD005461. CrossRef. Medline
6. Bedwell C, Levin K, Pett C, Lavender DT. A realist review of the partograph: when and how does it work for labour monitoring? *BMC Pregnancy Childbirth.* 2017;17(1):31. CrossRef. Medline
7. Ollerhead E, Osrin D. Barriers to and incentives for achieving partograph use in obstetric practice in low- and middle-income countries: a systematic review. *BMC Pregnancy Childbirth.* 2014;14:281. CrossRef. Medline
8. Sanghvi H, Mohan D, Litwin L, et al. Effectiveness of an electronic partogram: a mixed-method, quasi-experimental study among skilled birth attendants in Kenya. *Glob Health Sci Pract.* 2019;7(4):521–539. CrossRef
9. Litwin LE, Maly C, Khamis AR, et al. Use of an electronic Partograph: feasibility and acceptability study in Zanzibar, Tanzania. *BMC Pregnancy Childbirth.* 2018;18(1):147. CrossRef. Medline
10. Tadesse Y, Gelagay AA, Tilahun B, Endehabtu BF, Mekonnen ZA, Gashu KD. Willingness to Use Mobile based e-Partograph and Associated Factors among Care Providers in North Gondar Zone, Northwest Ethiopia. *Online J Public Health Inform.* 2019;11(2):e10. CrossRef. Medline
11. Rahman A, Begum T, Ashraf F, et al. Feasibility and effectiveness of electronic vs. paper partograph on improving birth outcomes: A prospective crossover study design. *PLoS One.* 2019;14(10):e0222314. CrossRef. Medline

Cite this article as: Hodgins S. A tablet-based tool for care during labor+attention to system requirements. *Glob Health Sci Pract.* 2019;7(4):498-499. https://doi.org/10.9745/GHSP-D-19-00384

GLOBAL HEALTH: SCIENCE AND PRACTICE
Dedicated to what works in global health programs
OPEN ACCESS

COMMENTARY

Using the Unmanned Aerial Vehicle Delivery Decision Tool to Consider Transporting Medical Supplies via Drone

Margaret Eichleay,[a] Emily Evens,[a] Kayla Stankevitz,[a] Caleb Parker[a]

We developed an unmanned aerial vehicle (UAV) Delivery Decision Tool to help health system decision makers identify their transport challenges and explore the potential utility and impact of UAVs on the broader health system.

INTRODUCTION

Unmanned aerial vehicles (UAVs) or drones are increasingly explored as a solution to transport challenges for medical goods, including emergency blood supplies, vaccines, medicines, diagnostic samples, and even organs, particularly for "last mile" delivery.[1,2] Proof-of-concept tests have demonstrated the technological viability of UAVs to safely transport medical supplies[3–6] and keep them within the required parameters for clinical viability.[7–9] A few research studies have determined the cost-effectiveness of adding UAVs to specific medical supply chains[1,10] or the optimal placement for UAV stations.[11–13] Yet very few cases of successful scale-up of UAVs for medical transport exist. Recent work has focused on proving a UAV can successfully complete a flight but has neglected to address health-system integration and long-term sustainability. As a result, decision makers lack information on how to explore the potential utility of UAVs in addressing medical transport challenges and, if determined useful, how to add UAVs to health systems or achieve use at scale.

Successful implementation of any health system change relies on a complex set of factors.[14] In the case of UAVs, not only will the selected technology need to be appropriate for the distance, weight, and size requirements of the transported items, but (1) it will need to operate within regulations, (2) the concept must be embraced by stakeholders, (3) financial resources need to be available, (4) human resources must be in place, and (5) operational procedures must be developed to effectively work within existing structures. When these aspects are not considered, operations can be delayed, inefficient, or fail entirely. In short, as with any intervention, the impact of UAVs on the broader health system should be considered before implementation.[15]

[a] FHI 360, Durham, NC, USA.
Correspondence to Margaret Eichleay (meichleay@fhi360.org).

Using the framework of the World Health Organization health system building blocks,[16] this commentary describes challenges for integrating UAVs into complex health systems and presents a tool for considering whether UAVs could help address medical transport challenges and how they can be integrated into health systems. Under the assumption that readers may be unfamiliar with UAVs, we begin with some general information (Box).

UAVS FOR MEDICAL TRANSPORT

Routine delivery of medical supplies via UAV is still relatively new. Until recently, there has been only 1 example of routine medical supply using UAV delivery. Since 2016, Zipline has operated drones for the Government of Rwanda, delivering up to 3 liters of blood within 30 minutes to health facilities that request it on demand.[17] In 2019, however, several other companies received approval to conduct routine flights. Matternet, a company that has conducted hundreds of test flights in Switzerland, routinely transports laboratory specimens within a North Carolina health system.[18] Alphabet's Wing will soon be delivering over-the-counter medications, via UAV in Canberra, Australia, after a year and a half of test deliveries.[19] Yet to date, none of these projects have published on the decision to implement UAV delivery, whether/how they integrated UAVs into existing health systems, or the impact the change has had on health care operations.

In addition to these larger-scale operations, many other pilot projects delivering medical goods are being implemented worldwide,[4,20,21] transporting items such as childhood vaccinations,[22,23] automatic external defibrillators,[11–13] snakebite antivenom,[24] tuberculosis sputum samples,[8,25] and sterile mosquitoes.[26] Although these projects are increasingly sharing information about operations, only a few resources for implementing UAV delivery projects exist [20,27,28] and none explore the feasibility and impact of adding this technology to the broader health system.

CONSIDERATIONS FOR INTRODUCING UAV TRANSPORT OF MEDICAL GOODS

Service Delivery

Adding any new product or process to a large existing health system can be challenging. Companies that have been transporting medical goods via drone for multiple years or are preparing to do so have yet to share information about how they integrate their processes with the health system, what processes have to change, how those impact workflows, or the impact they have on health outcomes. Understanding the potential risks and benefits of a systemic change is important during the consideration and implementation of a new technology for evaluating sustainability and the true cost of implementation. Whether the UAV system should be set up in parallel to or integrated into existing structures and systems is also under debate. As with many health innovations, there is concern that integration into existing structures takes too much time; yet, when parallel systems are devised, it fractures the health system, causing informational and operational silos that may result in future inefficiency.

Health Workforce

Having an adequate number of appropriately trained staff is a constant challenge in many low- and middle-income countries. Depending on the business model used to operate UAVs, adding them to the health system could have impacts on the health workforce. Health professionals may have a role in loading or unloading a UAV; confirming schedules; securing a loading, landing, or dropping site; documenting deliveries; launching the UAV; or instructing the UAV about its next location. Each of these actions would require training and time, so the volume of deliveries is important to consider as well. The tradeoffs between having health professionals perform these functions or hiring others to do it should be evaluated by decision makers before deciding on an operating model. This decision process is not different than typical health workforce considerations, but because the level of training and the steps required differ by UAV system, it is important to consider this information when designing or choosing any UAV delivery system.

Financing

The tradeoff between flight distance, the amount of weight a UAV can carry, and cost is another barrier to some applications of UAVs in health care. Applications for delivering medical goods in low- or middle-income countries via UAVs have been limited to very lightweight items (less than 5 kg) and relatively short distances (less than 50 km). However, the technology is constantly improving, such that new UAV models might carry 10 kg in weight up to 300 km in distance. The cost of these technologies is positively associated with distance and payload, and the value of UAV delivery to health provision has yet to be established.

Tools to explore feasibility and impact of UAVs in health systems do not yet exist.

Given the tradeoff between weight, distance, and cost, UAV transport will likely supplement medical supply chains, rather than replace road transport.[1] Understanding the conditions under which UAVs are cost-effective is a critical but complex area for investigation. Logistics management systems traditionally account for the costs of device operation and maintenance, as well as transport time, road condition, warehousing, and staff.[29] The cost of adding UAVs to the supply chain will be determined by these measures but unique considerations related to UAVs exist.

Understanding potential risks and benefits is important when considering and implementing a new technology.

First, the cost of the devices and the cost of operating and maintaining them vary substantially. Cost models will need to account for the specific UAV system used as well as staff training to operate and maintain them. Secondly, although road condition will not be a big factor in UAV operating costs, weather condition could be. Knowing the impact of wind, humidity, elevation, precipitation, and temperature on supply chain operations will be critical. In addition to direct costs of the technology and service provision, there are opportunity costs. Calculating the value of faster turnaround times for laboratory test results or the value of having a health provider remain in a facility for a day rather than transporting medical goods is difficult to measure but important for considering supplementing systems operations with UAVs. Although a cost-effectiveness modeling tool for UAV supply chains was created in 2019,[27] it does not yet incorporate the staff and

In addition to direct costs of the technology and service provision, opportunity costs should be considered.

BOX. What Is an Unmanned Aerial Vehicle (UAV)?[1,4]

- UAVs are aerial vehicles guided without an onboard crew.
- They can be piloted remotely or programmed to fly autonomously.
- A variety of models exist, each suited to different applications based on distance, payload, maneuverability, fuel sources, durability, need for landing, and other factors.

A quadcopter delivers an automated external defibrillator (AED).

opportunity costs. Full cost accounting methods are still needed.

Information Systems

Logistics information systems will be substantially impacted by the introduction of UAVs. If UAVs are added to an existing supply chain, all current documentation of what is stocked, transported, and received at multiple locations will have to change to account for method of transport (UAV or other). Because flights might originate from different locations than road route warehouses, these tracking systems will have to be aligned and combined to account for this complexity. Furthermore, logistics systems that optimize routing will require a new set of variables to determine the most efficient routing.

Other health information systems, such as patient and laboratory sample tracking systems, may also be impacted. Ideally, these would all be linked so that a physician who needs to send a patient's biologic sample to a testing laboratory could easily indicate that in the patient record, which would automatically trigger a request for UAV transport to the appropriate location, and the laboratory would know when it would be received. Although this level of systems integration remains unrealistic in many contexts, starting conversations about potential UAV integration and interoperability of systems early is key to efficient systems design.

Monitoring and evaluation systems are also needed. Supply chains are evaluated based on their ability to provide the correct quantity of goods, maintained within appropriate environmental conditions, to the correct location, on time, and at a competitive cost.[29] Although the indicators of success are unlikely to change with the introduction of UAVs, measuring how performance changes around the introduction of this new mode of transport will be key to understanding its value. A standardized approach to generate evidence around the introduction of UAVs into public health supply chains is being developed by the Interagency Supply Chain Group's Unmanned Aircraft Systems Coordinating Body and VillageReach.

Access to Essential Medicines

In the development sector, the goal of most UAV delivery programs will be improving access to health care, including essential medicines. Because access is such an important aspect of health systems operations, it is critical for UAV projects to measure the impact they have on access—not only by assessing how many more people are served, but also by ideally determining whether they are reaching those most in need, thereby measuring equity of access. Although difficult to measure until operations have been conducted for some time, planning for these evaluations early will produce higher-quality information.

Leadership and Governance

One priority governance issue for UAV use in general, not specific to delivery, is developing national and international regulations in the context of rapidly evolving technology. Air space is highly regulated by civil and international aviation authorities. Anticipating that commercial and federal use of UAVs will add aircraft to the airspace, regulators need to ensure that both large manned aircraft and smaller unmanned aircraft communicate with each other as well as authorities to ensure safe operations. Although UAVs fly at much lower altitudes than most manned aircraft, all aircraft pass through low-altitude space, thereby requiring coordination. The United Nations International Civil Aviation Organization, as well as many national civil aviation organizations, are tackling regulatory issues (e.g., how unmanned air traffic management systems will integrate with manned ones); developing methods to remotely identify UAVs and their operators; standardizing UAV regulations for the humanitarian and development sector; determining who is allowed to remove the threat of an unknown UAV; determining criteria for no-fly zones; and deciding under what conditions UAVs can fly beyond the operator's line of sight, over people, or

at night. Although these issues apply to all UAV flights, they are particularly important to consider before health systems can utilize UAVs for transport, as most applications require long-range transport, either start or end in urban locations, and in emergency situations may not be able to wait for dawn to take flight.

The complexity of health supply chains in low-resource settings requires dedicated stakeholder engagement. Often supply chains involve public, private, and nonprofit actors, national and international entities, siloed transport and tracking systems, and insufficient resources to strengthen them. However, UAV operations will be much more cost-effective if the technology can be used across silos and by multiple institutions.[1] Although using UAVs for public benefit is often met with enthusiastic response, stakeholders have demonstrated concerns about value for cost, privacy, security, how UAV regulations can be enforced,[30–32] and noise.[33] These concerns can pose real barriers to implementation. Recently, the Government of Ghana's plan to use UAV services to deliver blood products was opposed by the Ghana Medical Association, a stakeholder that was not sufficiently engaged before decision making and felt the investment misplaced.[34,35]

Many organizations have been working to ensure that decisions to use UAVs are made with as much information as possible. Over 5 years ago, a collaborative forum was developed to foster discussion and dissemination of UAV delivery work. The Unmanned Aerial Vehicles for Payload Delivery Working Group (UPDWG) currently has over 200 members from public, private, and nonprofit institutions. Recently, a more formal organizing body was created to strengthen coordination between stakeholders and share resources: Interagency Supply Chain Group Unmanned Aircraft Systems Coordinating Body.

As UAV technology has evolved, challenges have arisen in aligning technology partners with stakeholder needs. Ensuring applications of this technology address relevant problems by teams who understand both local contexts and the parameters of the technology remains a challenge. In lower- and middle-income countries, the use of UAVs involves multiple entities: companies who develop and implement the technology, international health and development organizations who implement programs, and national entities who manage the delivery of health care and commodities and implement regulations. We observed a need for a decision-support tool for implementers who are considering UAV technology for transporting medical goods that educates both operators and implementors on relevant considerations.

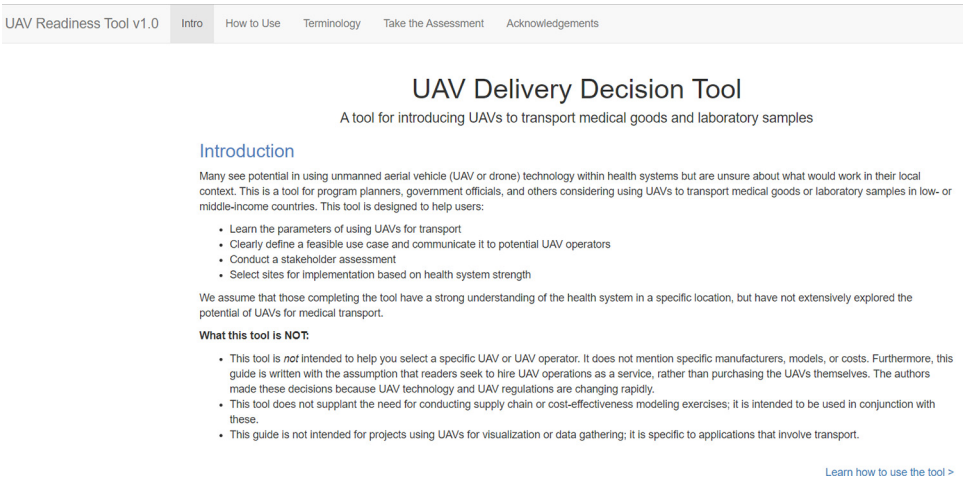
TOOL DEVELOPMENT PROCESS

The UAV Delivery Decision Tool was developed using 3 sources of information: a literature review, in-depth qualitative interviews, and expert review of the tool. A variety of readiness assessment tools exist in the literature, yet none of the existing tools offers a practical and directly applicable way to assess the introduction of UAV medical supply transport. Tools that assess organizational readiness for change abound and typically focus on motivation for change or capacity to change at a single institution,[14,36] whereas UAV medical transport will require collaboration between multiple institutions (e.g., health facilities, warehouses, regulatory bodies) in both the public and private sectors. Tools related to technology acceptance investigate why and when people use a new technology[37,38] and the potential to alter roles and responsibilities within existing structures.[39] But these tools typically assess existing, not novel, technologies. Drawing from the above review of considerations presented by the World Health Organization health system building blocks[16] and a review of organizational readiness tools,[14] we created a list of concepts applicable to the integration of UAVs into the health system and used it to inform the UAV Delivery Decision Tool.

Using UAVs for public benefit is often met with enthusiastic response, yet stakeholders have demonstrated concerns, including value for cost, privacy, and security.

Qualitative interviews were conducted in Nairobi and Turkana County, Kenya, in September 2018 with national- and county-level Ministry of Health officials, health facility staff, laboratory staff, medical supply transport management agencies, and a UAV operator in Kenya. All health personnel interviewed were part of the Afya Nyota ya Bonde project, a 5-year, U.S. Agency for International Development-funded project aiming to improve HIV care and treatment coverage, including commodity management and laboratory services. The project was interested in exploring the cost-effectiveness of using UAVs for transporting laboratory samples in Turkana County, Kenya, where current transport of dried blood spots for HIV viral load testing requires multi-stop transport over 350 km characterized by underdeveloped road networks, insecurity and violence, unreliable transportation, and flash flooding.[40–43] Participants expressed interest in the technology and its potential to save time, but they were also hesitant and wanted to understand more about costs and staffing. They mentioned the importance of sensitization, particularly given some of the political and financial

Ensuring applications of this technology address relevant problems by teams who understand both local contexts and the parameters of the technology remains a challenge.



Screenshot of UAV Delivery Decision Tool. © 2019/FHI 360

sensitivities involved in changing supply chains. Data were also collected on the current costs and frequency of ground-based transport (presented in a separate paper). Information from the literature review and these interviews were synthesized to produce a decision-support tool.

We requested feedback on the first draft of the tool from members of UPDWG, program directors, and government officials. Most of the feedback received was from UPDWG and was incorporated into the present version.

TOOL DESCRIPTION

The UAV Delivery Decision Tool (https://fhi360.shinyapps.io/UAVDeliveryDecisionTool/) is a 4-part tool developed in Shiny, version 1.1.0,[44] an open-access application development package. In Part 1, users define the transport problem they are trying to solve using UAVs and its root cause to determine whether UAVs are a potential solution. Part 2 includes questions regarding transport parameters (origin, destination, distance), the transported goods (weight, dimensions, temperature), and geographic context (terrain, security). Responses to Parts 1 and 2 produce an editable, user-tailored document offering guidance and a clearly defined use case. Parts 3 and 4 contain instructions for completing offline worksheets to identify and analyze stakeholders and to select preliminary sites for conducting UAV operations. Various scenarios can be tested by completing the tool repeatedly and adjusting inputs. We encourage those using the tool to share what they learn from it and from any UAV transport operations through UPDWG (updwg.org) to further encourage collaboration and learning.

THE WAY FORWARD

Interest in using UAVs to transport medical goods is currently high, but the health sector lacks structured guidance to systematically consider the feasibility, utility, and impact of using UAV transport and evidence regarding implementation. The UAV Delivery Decision Tool is designed to help implementers consider their options and UAV developers to understand the context within which their products need to operate. If a decision to implement UAV delivery is made, gathering evidence on those activities will enable sustainable and careful integration of this new technology that is likely to revolutionize the transport sector in the next decade.

Acknowledgments: Without the programming expertise of James MacCarthy, this tool would not have been possible, and for that the authors are deeply grateful. We also wish to thank David Onyango and Michael Loogos, who facilitated the field work portion of this project. For sharing their thoughts and experience, we would like to thank our key informants in Turkana County and Nairobi, Kenya. We are also grateful for the input provided by members of the UPDWG community.

Funding: This work was supported by FHI 360's Research and Evaluation Strategic Initiative.

Competing interests: None declared.

REFERENCES

1. Wright C, Rupani S, Nichols K, Chandani Y, Machagge M. *What Should You Deliver by Unmanned Aerial Systems? The Role of Geography, Product, and UAS Type in Prioritizing*

Deliveries by UAS. Arlington, VA: JSI Research & Training Institute, Inc.; 2018.

2. Zraick K. Like 'Uber for Organs': Drone delivers kidney to Maryland woman. New York Times. April 30, 2019. https://www.nytimes.com/2019/04/30/health/drone-delivers-kidney.html. Accessed June 4, 2019.
3. USAID Global Health Supply Chain Program-Procurement and Supply Management. Unmanned Aerial Vehicles Landscape Analysis: Applications in the Development Context. Washington, DC: Chemonics International Inc.; 2017.
4. Swiss Foundation for Mine Action (FSD). *Drones in Humanitarian Action: A Guide to the Use of Airborne Systems in Humanitarian Crises*. Geneva: FSD; 2016.
5. Mendelow B, Muir P, Boshielo BT, Robertson J. Development of e-Juba, a preliminary proof of concept unmanned aerial vehicle designed to facilitate the transportation of microbiological test samples from remote rural clinics to National Health Laboratory Service laboratories. *S Afr Med J*. 2007;97(11 Pt 3):1215–1218. Medline
6. Amukele T, Ness PM, Tobian AA, Boyd J, Street J. Drone transportation of blood products. *Transfusion*. 2017;57(3):582–588. CrossRef. Medline
7. Amukele TK, Hernandez J, Snozek CLH, et al. Drone transport of chemistry and hematology samples over long distances. *Am J Clin Pathol*. 2017;148(5):427–435. CrossRef. Medline
8. Amukele TK, Street J, Carroll K, Miller H, Zhang SX. Drone transport of microbes in blood and sputum laboratory specimens. *J Clin Microbiol*. 2016;54(10):2622–2625. CrossRef. Medline
9. Amukele TK, Sokoll LJ, Pepper D, Howard DP, Street J. Can unmanned aerial systems (drones) be used for the routine transport of chemistry, hematology, and coagulation laboratory specimens? *PLoS One*. 2015;10(7):e0134020. CrossRef. Medline
10. Haidari LA, Brown ST, Ferguson M, et al. The economic and operational value of using drones to transport vaccines. *Vaccine*. 2016;34(34):4062–4067. CrossRef. Medline
11. Pulver A, Wei R, Mann C. Locating AED enabled medical drones to enhance cardiac arrest response times. *Prehosp Emerg Care*. 2016;20(3):378–389. CrossRef. Medline
12. Claesson A, Fredman D, Svensson L, et al. Unmanned aerial vehicles (drones) in out-of-hospital-cardiac-arrest. *Scand J Trauma Resusc Emerg Med*. 2016;24(1):124. CrossRef. Medline
13. Boutilier JJ, Brooks SC, Janmohamed A, et al; Rescu Epistry Investigators. Optimizing a drone network to deliver automated external defibrillators. *Circulation*. 2017;135(25):2454–2465. CrossRef. Medline
14. Dearing JW. Organizational readiness tools for global health intervention: a review. *Front Public Health*. 2018;6(56):56. CrossRef. Medline
15. Swanson RC, Mosley H, Sanders D, et al. Call for global health-systems impact assessments. *Lancet*. 2009;374(9688):433–435. CrossRef. Medline
16. World Health Organization (WHO). *Everybody's business: strengthening health systems to improve health outcomes: WHO's framework for action*. Geneva: WHO; 2007.
17. Our Impact. Zipline website. https://flyzipline.com/impact/. Accessed October 4, 2019.
18. UPS partners with Matternet to transport medical samples via drone across hospital system in Raleigh, N.C. [press release] Atlanta, GA: UPS; March 26, 2019.
19. Porter J. Google's Wing drones approved to make public deliveries in Australia. *The Verge*. April 9, 2019. https://www.theverge.com/2019/4/9/18301782/wing-drone-delivery-google-alphabet-canberra-australia-public-launch. Accessed October 4, 2019.
20. United States Agency for International Development (USAID). *UAVs in Global Health: Defining a Collective Path Forward*. Washington, DC: USAID; 2018.
21. Wingcopter, DHL and GIZ pilot drone delivery in Tanzania. Wingcopter website. https://wingcopter.com/deliverfuture/. Accessed November 4, 2019.
22. Held A. Drones deliver vaccines on island nation of Vanuatu. National Public Radio website. https://www.npr.org/sections/goatsandsoda/2018/12/21/678101149/drones-deliver-vaccines-on-island-nation-of-vanuatu. Published December 21, 2018. Accessed October 4, 2019.
23. Thorne J. Drones swoop into the Congo to strengthen the country's vaccine supply chain. *Geekwire*. https://www.geekwire.com/2019/drones-swoop-congo-strengthen-countrys-vaccine-supply-chain/. Published August 8, 2019. Accessed November 4, 2019.
24. Meier P, Bergelund J. *Field-testing the First Cargo Drone Deliveries in the Amazon Rainforest*. https://blog.werobotics.org/wp-content/uploads/2017/02/WeRobotics-Amazon-Rainforest-Cargo-Drones-Report.pdf. WeRobotics; 2017.
25. Meier P, Soesilo D. Case Study No. 2: *Delivery - Using Drones for Medical Payload Delivery in Papua New Guinea*. https://drones.fsd.ch/wp-content/uploads/2016/04/Case-Study-No2-PapuaNewGuinea.pdf. Geneva, Switzerland; Swiss Foundation for Mine Action (FSD); 2014.
26. Wu K. Do not fear the drones air-dropping 50,000 mosquitoes from above. *Smithsonian Magazine*. August 20, 2018. https://www.smithsonianmag.com/science-nature/do-not-fear-drones-air-dropping-50000-mosquitoes-above-180970068/. Accessed October 4, 2019.
27. JSI Research and Training Institute, Llamasoft. *What Should You Deliver by Autonomous Aerial Systems? Tool for Determining Cost Effective Use Cases for AAVs*. https://www.updwg.org/resource-library/. Published January 2019. Accessed Sep 30, 2019.
28. United States Agency for International Development (USAID), Global Health Supply Chain (GHSC) Program. *Unmanned Aerial Vehicle Procurement Guide: Specifications, questions and other criteria to consider*. Washington, DC: USAID; 2018.
29. United States Agency for International Development (USAID), DELIVER PROJECT, Task Order 1. *The Logistics Handbook: A Practical Guide for the Supply Chain Management of Health Commodities*. Arlington, Va.: USAID | DELIVER PROJECT, Task Order 1; 2011.
30. Eichleay M, Mercer S, Murashani J, Evens E. *Using Unmanned Aerial Vehicles for Development: Perspectives from Citizens and Government Officials in Tanzania*. Durham, NC: FHI 360; 2016.
31. Eyerman J, Hinkle K, Letterman C, et al. *Unmanned Aircraft and the Human Element: Public Perceptions and First Responder Concerns*. Research Triangle Park, NC: Institute for Homeland Security Solutions; 2013.
32. Fraser A. Witchcraft and explosions: perceptions of drones in Africa. WeRobotics Blog. https://blog.werobotics.org/2017/08/10/witchcraft-and-explosions-perceptions-to-drones-in-africa/. Published August 10, 2017. Accessed October 4, 2019.
33. Wiggers K. Customers compare the noise from Alphabet spinout Wing's delivery drones to a chainsaw. *VentureBeat*. Published December 27, 2018. Accessed October 4, 2019.
34. Ghana drones: Row over blood-delivery devices. BBC website. https://www.bbc.com/news/world-africa-46543442. Published December 12, 2018. Accessed October 4, 2019.
35. Novelty is NOT Innovation - the Story of Fly Zipline Ghana. IMANI Center for Policy & Education website. https://imaniafrica.org/2018/12/06/imani-alert-novelty-is-not-innovation-the-story-of-fly-zipline-ghana/. Published December 6, 2018. Accessed October 4, 2019.

36. Weiner BJ, Amick H, Lee SY. Conceptualization and measurement of organizational readiness for change: a review of the literature in health services research and other fields. *Med Care Res Rev.* 2008;65(4):379–436. CrossRef. Medline

37. Parasuraman A. Technology Readiness Index (Tri). *J Serv Res.* 2000;2(4):307–320. CrossRef

38. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. *Manage Sci.* 1989;35(8):982–1003. CrossRef

39. Edmondson AC, Bohmer RM, Pisano GP. Disrupted Routines: Team Learning and New Technology Implementation in Hospitals. *Adm Sci Q.* 2001;46(4):685–716. CrossRef

40. Hicks D, Raja S, Heinen F. *Kenya- Improving Health Systems: Public Sector Healthcare Supply Chain Strategic Network design for KEMSA.* Washington, DC: The World Bank; 2009.

41. Johannes EM, Zulu LC, Kalipeni E. Oil discovery in Turkana County, Kenya: a source of conflict or development? *African Geographical Review.* 2015;34(2):142–164. CrossRef

42. Nanok J. *Turkana County Investment Plan* 2016-2020. Kenya: UNDP; 2015.

43. World Health Organization. *Kenya Service Availability and Readiness Assessment Mapping: SARAM Report.* Nairobi, Kenya: Government of Kenya; 2014.

44. shiny: Web Application Framework for R. R package version 1.1.0. [computer program]. Version 1.1.02018. R Studio.

Peer Reviewed

Received: April 2, 2019; **Accepted:** October 15, 2019; **First Published Online:** December 4, 2019

Cite this article as: Eichleay M, Evens E, Stankevitz K, Parker C. Using the unmanned aerial vehicle delivery decision tool to consider transporting medical supplies via drone. *Glob Health Sci Pract.* 2019;7(4):500-506. https://doi.org/10.9745/GHSP-D-19-00119

GLOBAL HEALTH: SCIENCE AND PRACTICE
Dedicated to what works in global health programs
OPEN ACCESS

COMMENTARY

Using Digital Technology for Sexual and Reproductive Health: Are Programs Adequately Considering Risk?

Loraine J. Bacchus,[a] Kate Reiss,[a] Kathryn Church,[b] Manuela Colombini,[a] Erin Pearson,[c] Ruchira Naved,[d] Chris Smith,[a,e] Kathryn Andersen,[c] Caroline Free[a]

Digital technologies provide opportunities for advancing sexual and reproductive health and services but also present potential risks. We propose 4 steps to reducing potential harms: (1) consider potential harms during intervention design, (2) mitigate or minimize potential harms during the design phase, (3) measure adverse outcomes during implementation, and (4) plan how to support those reporting adverse outcomes.

INTRODUCTION

Health care is increasingly being delivered through digital channels such as the internet, mobile phone messaging, social media, apps, voice, video messaging, and telemedicine. This trend has been facilitated by diffusion of mobile technology and rapid advances in artificial intelligence. Digital communication channels offer wide coverage, allow messaging to be targeted to particular groups or individuals, and offer potential for enhancing the delivery of sexual and reproductive health and rights (SRHR) information and support.

Recent developments in the SRHR field include the provision of online testing for sexually transmitted infections (STIs) that have been shown to almost double the uptake of STI tests[1] and e-contraception whereby the oral contraceptive pill can be ordered online.[2] Telemedicine in sexual and reproductive health (SRH) can overcome geographic or social and behavioral barriers to accessing services and facilitate self-use of products or services.[3] It has been used to support medication abortion and facilitate distribution of abortifacient pills backed up by remote care and support.[4] Interventions targeting a range of populations and SRHR topics across different cultural contexts have been shown to be acceptable to the end user and feasible to implement.[4–8] Interventions can be designed to be accessible across socioeconomic groups and to those at high risk.[2,9,10] Improvements in knowledge and contraceptive or health-seeking behavior have been demonstrated.[6,11–13] However, not all studies show benefits, as exemplified by the Reiss et al. intervention in Bangladesh that showed no effect on contraceptive use.[14]

Sensitivity of SRH and SRH Services

SRH and SRH services are highly sensitive. There can be considerable social disapproval of sexual behaviors within some groups, such as adolescents and men who have sex with men, or outside of marriage. HIV and other STIs are stigmatized and decisions about fertility are highly influenced by partners as well as by members of the broader family and community. Failure to keep SRH service use confidential or disclosure of HIV/STI status can result in conflict with or loss of support of parents, stigma, blame, discrimination, or new or escalating verbal or physical violence.[15–17] Women in many settings, in particular those living in more patriarchal, socially conservative contexts, are also often victims of reproductive coercion, through which their autonomy over reproductive choice is greatly impaired.[18–19] Women's requests to use contraception can result in suspicions of infidelity, and many women fear violence if they request to control their own fertility.[20] In this context, a mechanism by which harm may occur to women is through breaches of confidentiality. If a partner or family member becomes aware of the woman's access to SRH services, they may restrict the woman from returning to a health facility and accessing health care.

Reproductive coercion is a form of intimate partner violence (IPV), which includes physical, sexual, and emotional abuse and controlling behaviors by a current or former partner.[21] IPV is one of the most common forms of violence against women with 1 in 3 women globally having experienced it in their lifetime.[22] The estimated prevalence of reproductive coercion among

[a] London School of Hygiene & Tropical Medicine, London, England.
[b] Marie Stopes International, London, England.
[c] Ipas, Chapel Hill, NC, USA.
[d] icddr,b, Dhaka, Bangladesh.
[e] School of Tropical Medicine and Global Health, Nagasaki University, Nagasaki, Japan.
Correspondence to Loraine Bacchus (Loraine.Bacchus@lshtm.ac.uk).

women experiencing IPV is 8% to 16% in U.S. studies.[23–24] Negative SRH outcomes, such as unintended pregnancy and termination of pregnancy, are potential consequences of reproductive coercion.[25–28] A study in Bangladesh found that IPV was associated with women coming to the health facility alone for menstrual regulation rather than being accompanied by their husband or partner.[29]

These issues have rendered privacy and confidentiality a key tenet of quality SRH information and services. They are especially important for protecting vulnerable women and those experiencing IPV.

Potential Risks of Digital SRH Information Support and Services

Although the extensive use of digital technologies affords opportunities for SRH promotion and service provision, interventions that reach into users' homes and personal spaces may also entail particular risk. For example, interventions designed to increase adherence to antiretroviral medication for patients who have HIV may inadvertently result in disclosure of HIV status. Interventions promoting contraception may inadvertently disclose contraceptive use that others may not condone.[14]

Mobile phones have features that can be used to enhance privacy. In SRH contexts, sensitive personal content delivered by digital media can be confidential, especially if phone features such as passwords are not shared and alerts are switched off.[9] However, boundaries in relation to mobile phone access and concepts of privacy vary between individuals and contexts. Phones may be shared within families, hindering the potential to deliver sensitive personal information.[30] Women in abusive situations may have their digital media use controlled by their partner and family members, or digital media may be used to perpetrate abuse (i.e., harass, stalk, and monitor).[31]

BOX. Steps to Reduce Potential Harms in Providing Sexual and Reproductive Health Services

1. Consider potential harms during intervention design. With input from users and key stakeholders:
   - Develop theoretical frameworks to elaborate mechanisms through which adverse outcomes may occur.
   - Understand how digital media are used, shared, and kept private.
   - Understand the sensitivity, stigma, and social and power dynamics influencing SRH in the context for which the intervention is planned.
2. Mitigate or minimize potential harms in the design phase.
   - Consult with potential users in intervention development to determine whether privacy can be achieved. Consider the mode of delivery and intervention features required to afford privacy and confidentiality, when wanted.
   - Test and refine interventions with input from users receiving the intervention as planned.
3. Measure adverse outcomes.
   - Use research methods to reduce reporting bias such as by using standardized, validated measures.
   - Follow ethical guidelines for conducting research on violence.
4. Plan how to support those reporting adverse events, including IPV.
   - Provide links to existing services and/or staff training according to setting.

MINIMIZING AND RESPONDING TO RISKS OF HARM IN SRH DIGITAL INTERVENTIONS

As with other forms of communication and service delivery in SRHR, potential harms of using digital technology must be first considered, then mitigated or minimized in the intervention design, and measured during evaluation or implementation. Those who deliver interventions should also plan how to respond to reports of harm. We discuss 4 key steps that should be taken to minimize and respond to risk of harm in SRHR interventions, specifically focusing on digital interventions (Box).

1. Consider Potential Harms

It is common practice to develop theories to explain how intended effects of interventions may occur and to explore these in evaluations. This practice can inform understanding as to why interventions did or did not work, what the active components were, and how they may be transferred to other contexts or populations. In addition to identifying the mechanisms behind intended benefits, it is also important to identify "dark logic" models[32] (i.e., the causal pathways from interventions to potential adverse outcomes). Mechanisms for surfacing potential harms include consulting key stakeholders at an early stage of intervention development who have deep insight into how interventions work in local contexts. Other mechanisms include building a comparative understanding of potential harm across similar interventions through a review of the evidence (where such an evidence base exists) and reflecting on how broader sociocultural, political, and economic forces may constrain recipients

of the intervention (providers, clients) and cause unintended negative consequences.[32] For digital intervention design, key stakeholder interviews must include interviews with potential users to understand how digital media are used and shared and whether they could be kept private. It is also important to understand the sensitivity, stigma, social, and power dynamics influencing SRH in the intervention context. Given the infrequency of many negative outcomes, it may also be important to consider the use of meta-analysis to understand the existing evidence about potential risk, since individual studies may remain underpowered to detect effects.[32–33]

2. Mitigate and Minimize Harms

In conjunction with users, those designing interventions should consider which risk reduction strategies are needed. For interventions delivered by digital media, the first consideration is the choice of media for communication. Designers should consider the balance between the convenience of the digital intervention, maintenance of privacy and confidentiality, cultural acceptability, and effectiveness of the intervention. These factors will differ between contexts. For example, the Reiss et al. intervention in Bangladesh used voice messages[14] because participants in their formative research indicated that they felt voice messages were more private as they do not remain on the phone.[34] However, for some, voice messages may be more intrusive than text messages or other digital media because recipients have to answer the call to receive the message irrespective of where they are, what they are doing, or who they are with. In some contexts, use of text messages is more common than voice calls, thereby arousing less suspicion.

Once designers select the digital media, they can further minimize risk by using a range of options.[35] Where privacy cannot be achieved, designers may opt to send content specifically designed for sharing. For example, in work that we conducted in Cambodia[36] targeting factory workers, the sensitivity of sexual activity among unmarried female factory workers and shared living arrangements led to the intervention promoting contraceptive use and services via social media videos designed to be viewed alone or shared with groups of friends. Alternatively, designers may opt to send general content rather than personalized information that may reveal confidential behaviors. The use of general content has been shown to be both effective in changing attitudes and safe in 3 different settings.[37–38]

Further strategies to prevent harm in the case of content being viewed by others include careful naming of apps or wording of notifications, using discreet icons, avoiding stating the source within each message, and using general or untraceable telephone numbers rather than a short code.[9] Strategies to facilitate privacy include using passwords, blocking alerts, and using protective firewalls for sensitive applications. Other features that can protect recipients from having content viewed by others include escape buttons on websites and apps that allow pages to be shut quickly or switched to other sites[39] and customizable privacy settings including for data storage.[9] For example, interventions may be designed with data-purging mechanisms and without unnecessary automatic data collection systems that are often included as standard features.[35] Recipients can also be advised to delete messages or search histories if they have concerns about privacy.[9] Timing of content delivery can be critical to ensure messages are only sent when privacy can be assured. In a messaging intervention for sex workers, potential recipients only wanted to receive messages on Saturday mornings when they were not working.[40] In Bolivia, Palestine, Tajikistan, and the UK, where people have opted to receive push content (written messages sent to their phones), recipients reported them to be highly acceptable, safe, and effective in changing attitudes, intentions, and some preventative behaviors.[9,38,41,42] In other circumstances or contexts it may only be considered appropriate to use pull content (only sending messages or content on request) to reduce the potential for harm or it may be necessary to avoid outbound messages altogether.[43]

After considering the above factors early in the design phase, designers should test intervention prototypes and content with users. This phase should involve delivering the intervention in real time as intended and seeking feedback about recipients' experiences of the intervention, its acceptability, intrusiveness, ability to keep content private when wanted, and perceived impacts on the knowledge, attitudes, and behaviors targeted. In our previous interventions, many recipients wanted and chose to share content with siblings, partners, or mothers, but they felt it was important to have the ability to keep content private when they chose.[8,38,41,42]

3. Measure Potential Harms

Measuring potential risks is key to establishing the success of steps taken to mitigate or minimize

harms. Research offers critical opportunities to assess potential harms. Yet adverse event measurement is often limited to clinical research studies and remains rare and inconsistent in social and behavioral intervention evaluation.[32,33,44,45] Despite the known association between SRH or SRH service use and stigma, blame, discrimination, new or escalating verbal or physical violence from partners, family, or others (including violence during pregnancy, abortion, or post-HIV diagnosis),[46–49] it is surprising how few SRHR intervention trials have evaluated the impact of interventions on potential harms. For example, of the 5 trials included in a systemic review of mobile phone-based interventions for improving contraceptive use, only 1 assessed unintended adverse outcomes.[11] In a systematic review of STI/HIV partner notification, 5 of 26 studies assessed the number of harmful events reported.[50] Nonetheless, where evaluated, rare but important harms have been recorded such as perceived breaches of confidentiality regarding previously undisclosed HIV status.[51]

The importance of measurement is illustrated by the Reiss et al. intervention.[14] In the intervention development, the researchers considered the potential for harm among a group of women seeking abortion in a context where partner violence is common, and they took a number of steps to mitigate harm.[14] The research involved a substantive formative phase to develop and define the intervention. This informed the authors' understanding of user needs during the post-menstrual regulation period and the acceptability of having information delivered on this topic via mobile phone.[14,34] During in-depth interviews, user inputs were sought on the message content, mode of delivery, and issues related to confidentiality and privacy where prototype messages were played and shown to participants. Concerns were not raised at that stage by participants.[34] User-testing involved sharing draft content with contraceptive users.[14] During the trial recruitment process, researchers played an example of message content to all participants to check that they were comfortable receiving intervention voice messages.[14] The subsequent evidencing of the IPV findings in the Reiss randomized controlled trial demonstrates that even in-depth formative work and screening may be insufficient to remove risk.[14] This may reflect the challenges in overcoming social desirability reporting bias or power imbalances, or the challenges of asking users about hypothetical untested scenarios during the development phase. A pilot test of the intervention delivered as planned, followed by in-depth interviews, may have been more likely to reveal harms associated with the intervention, although the sample sizes of a pilot still would not have determined the degree of risk. However, measurement in the randomized controlled trial was key in preventing the scale-up of an intervention with no benefits, but with risk of harm. It is essential that the unintended consequences of all SRHR interventions, digital or otherwise, be consistently collected and reported on.

Challenges to Measuring Intimate Partner Violence

Given the documented links described earlier between abuse (including IPV) and SRH outcomes, reproductive coercion and other forms of abuse should be measured consistently in SRHR interventions. The use of valid and reliable measures will permit comparative analysis between settings and types of intervention. Although valid and reliable measures of IPV exist, the task of measuring IPV is complicated. Studies of IPV commonly use multicomponent self-report measures describing specific behaviors, but many validated measures are too long to be included in surveys, particularly if IPV is not the focus. Psychological and emotional abuse are particularly challenging to measure.[52–54] Due to the insidious nature of coercive and controlling behavior, women may not label them as abusive or misinterpret survey items leading to underreporting.[55] They may also be reluctant to identify with some forms of abuse. For example, in the context of a relationship where there has been prior consensual sex, a woman may be more hesitant to label an event as coerced or forced sex, even if consent was not given.[56] However, even physical violence, the most tangible form of violence, is difficult to measure in different contexts and tools vary in their comprehensiveness, affecting calculated prevalence rates.[53]

Advances in overcoming reporting bias, such as the collection of contextual data and outcomes (e.g., injury or other impacts),[54,57,58] may be ethically and practically challenging for wider SRHR evaluation where IPV measurement is not the primary objective. Lengthy surveys may compromise broader data quality by reducing response rates.[59] It is advisable to augment quantitative evaluation of IPV with qualitative methods to enhance interpretation and improve the reliability and validity of the findings, as well as explore unanticipated harmful outcomes post-intervention.[60–61] Clearly, there is a need to find balance between including questions on key dimensions

of IPV while being mindful of these ethical or practical concerns. The Bangladesh trial used closed-answer questions about specific acts of violence and followed the quantitative phase with in-depth interviews exploring the violence outcome.[14] The World Health Organization developed ethical and safety guidelines for researching violence against women, given the inherent risks to participants and researchers.[62,63] Recommendations include researcher training (i.e., for interviewing, using safety protocols, and dealing with participant distress and requests for help), provision to participants of up-to-date referral information for good-quality local organizations offering support to women experiencing violence,[63] and provision of psychosocial support for researchers experiencing vicarious trauma.[64] However, a pragmatic approach may be necessary as implementation of the guidelines requires additional resources in terms of time, finances, staffing, and technical support. This may include using a combination of remote and in-country research capacity strengthening sessions with local researchers. An advisory group with expertise in gender-based violence and intervention research can review safety and distress protocols for researchers and participants, as well as offer guidance when potential issues of harm and adverse events arise as a result of the research. Community-based organizations offering support should be identified, alerted to the intervention, and prepared for the potential increase in referrals as a result of the study.

4. Plan How to Support Women Experiencing IPV

Although digital interventions pose potential risks, they also provide significant opportunities for women experiencing IPV. For example, in clinical settings that lack privacy and where face-to-face communication can be overheard, the use of technology can reduce the risk of harm resulting from breaches of confidentiality. For some individuals, such interventions may be the only feasible way to obtain information and services. SRH services offer an entry point where women can safely disclose IPV experiences to their health care provider and access support and referrals within and outside of the health system.[65–67] Technology may also offer a greater sense of anonymity, which can reduce the anticipated stigma associated with direct disclosure to a health care provider.[68] The potential also exists to identify women experiencing IPV through digital SRHR outreach (information, mobile services) and streamline their referral for counseling and support. Evidence-based global guidelines on management of IPV state that health care providers should receive training in how to identify clients affected by IPV and provide first-line support that includes empathic listening, psychosocial support, and referral to appropriate services.[69] However, at present, evidence of effective interventions for addressing IPV in SRH is limited, particularly for low- and middle-income countries, and focus mainly on antenatal care settings.[69] Similarly, there has been little investigation into whether digital SRHR interventions can be effective in providing support for individuals who are experiencing IPV. This support has been requested during development of a text message intervention designed to reduce unintended pregnancy among female sex workers in Kenya, where many participants expressed the need for a service to report violence and receive emergency assistance. The intervention was adapted to incorporate an automatic response to any direct message into the system and information on gender-based violence support services.[40]

In recent years, there has been a proliferation of web-based apps that provide information about IPV, for example, on services and healthy relationships, as well as interactive tools for assessing risk and developing safety plans.[70–72] These apps have primarily targeted women in the general population, with the exception of 1 study, a nurse home visitation program for pregnant and postnatal women using mHealth.[68]

Additionally, the evidence is skewed toward high-income countries (i.e., Australia, Canada, New Zealand, and the United States), and little is known about how well these technologies work in low- and middle-income settings where there are greater structural barriers to women seeking help. Also, it is not clear how these apps may be adapted for use in SRH settings. Jewkes et al. (2019) highlights the need for more formative research and evaluation of their efficacy before roll out.[73] This research should include qualitative exploration of women's views, particularly their comfort in using apps to access information and support and their preference in engaging with apps alone versus with a health care provider. Research on health care providers' views should also include factors they should consider when trying to enroll women who are experiencing IPV into studies using apps. Cognitive interviewing techniques can be used to improve content, understandability, and design features. Finally, pilot and proof-of-concept studies are needed before progressing to trials or observational studies.

Whether addressing violence is the primary objective of an intervention or a component in a broader SRH intervention, the potential for this harm needs to be considered. In violent relationships, where abusive partners may search mobile phones and browsing histories, accessing support for violence or SRHR interventions could trigger a violent episode.[73] The need to protect women and children, ensure their safety, and avoid the risk of retaliation while they seek help is critical. Further research is needed to understand the relative acceptability, risks, and effectiveness of different models.

CONCLUSION

As mobile phone networks proliferate throughout low- and middle-income countries, digital technologies offer huge potential to support women to achieve positive SRHR outcomes. Technology can make information and services available when and where they are needed, and can facilitate a broader shift toward user-controlled products and services, including for family planning. However, delivering SRHR services and offering information and support to empower women to take control of their health and fertility may, in some instances, pit individuals' sexual health rights and reproductive autonomy against conservative or patriarchal social norms; this conflict may in turn place some women at risk. Nevertheless, digital technologies have potential to do good if they are well designed and implemented because they are highly accessible and can allow scale to be achieved at low cost.

We recommend that program designers follow 4 steps to reduce potential harms (Box 1). Taking these steps to develop interventions that are safe by design is key to the proliferation of digital SRH information, support, and services.

Acknowledgments: LJB, CF, and KR wrote the commentary with comments from all other authors.

Competing interests: None declared.

REFERENCES

1. Wilson E, Free C, Morris TP, et al. Internet-accessed sexually transmitted infection (e-STI) testing and results service: A randomized, single-blind, controlled trial. *PLoS Med*. 2017;14(12):e1002479. CrossRef. Medline
2. Rezel E, Free C, Syred J, Baraister P. Online Contraception - Innovation to Expand Choice in Access. Poster presented at: Faculty of Sexual & Reproductive Healthcare of the Royal College of Obstetricians & Gynaecologists Annual Scientific Meeting; 27–28 April 2017; Cardiff, Wales.
3. Remme M, Narasimhan M, Wilson D, et al. Self care interventions for sexual and reproductive health and rights: costs, benefits, and financing. *BMJ*. 2019;365:l1228. CrossRef. Medline
4. Endler M, Lavelanet A, Cleeve A, Ganatra B, Gomperts R, Gemzell-Danielsson K. Telemedicine for medical abortion: a systematic review. *BJOG*. 2019;126(9):1094–1102. CrossRef. Medline
5. L'Engle KL, Mangone ER, Parcesepe AM, Agarwal S, Ippoliti NB. Mobile phone interventions for adolescent sexual and reproductive health: A systematic review. *Pediatrics*. 2016;138(3):e20160884. CrossRef. Medline
6. Daher J, Vijh R, Linthwaite B, et al. Do digital innovations for HIV and sexually transmitted infections work? Results from a systematic review (1996-2017). *BMJ Open*. 2017;7(11):e017604. CrossRef. Medline
7. French RS, McCarthy O, Baraitser P, Wellings K, Bailey JV, Free C. Young people's views and experiences of a mobile phone texting intervention to promote safer sex behavior. *JMIR mHealth Uhealth*. 2016;4(2):e26. CrossRef. Medline
8. Smith C, Ly S, Uk V, Warnock R, Free C. Women's views and experiences of a mobile phone-based intervention to support post-abortion contraception in Cambodia. *Reprod Health*. 2017; 14(1):72. CrossRef. Medline
9. Free C, McCarthy O, French RS, et al. Can text messages increase safer sex behaviours in young people? Intervention development and pilot randomized controlled trial. *Health Technol Assess*. 2016;20(57):1–82. CrossRef. Medline
10. Ampt FH, Mudogo C, Gichangi P, et al. WHISPER or SHOUT study: protocol of a cluster-randomised controlled trial assessing mHealth sexual reproductive health and nutrition interventions among female sex workers in Mombasa, Kenya. *BMJ Open*. 2017;7(8):e017388. CrossRef. Medline
11. Smith C, Gold J, Ngo TD, Sumpter C, Free C. Mobile phone-based interventions for improving contraception use. *Cochrane Database Syst Rev*. 2015;(6):CD011159. CrossRef. Medline
12. Sondaal SFV, Linda Browne J, Amoakoh-Coleman M, et al. Assessing the effect of mHealth interventions in improving maternal and neonatal care in low-and middle-income countries: a systematic review. *PLoS One*. 2016;11(5). CrossRef. Medline
13. Burns K, Keating P, Free C. A systematic review of randomised control trials of sexual health interventions delivered by mobile technologies. *BMC Public Health*. 2016;16(1):778. CrossRef. Medline
14. Reiss K, Andersen K, Pearson E, et al. Unintended Consequences of mHealth Interactive Voice Messages Targeting Post-Menstrual Regulation Contraceptive Use in Bangladesh: Intimate Partner Violence Results From a Randomized Controlled Trial. *Glob Heal Sci Pract*. 2019;7(3):386–403. CrossRef. Medline
15. Naigino R, Makumbi F, Mukose A, et al. HIV status disclosure and associated outcomes among pregnant women enrolled in antiretroviral therapy in Uganda: a mixed methods study. *Reprod Health*. 2017;14(1). CrossRef. Medline
16. Duncan B, Hart G, Scoular A, Bigrigg A. Qualitative analysis of psychosocial impact of diagnosis of *Chlamydia trachomatis*: implications for screening. *BMJ*. 2001;322(7280):195–199. CrossRef. Medline
17. Nack A. Damaged goods: women managing the stigma of STDs. *Deviant Behav*. 2000;21(2):95–121. CrossRef
18. Grace KT, Anderson JC. Reproductive coercion: a systematic review. *Trauma Violence Abuse*. 2018;19(4):371–390. CrossRef. Medline
19. Fay K, Yee L. Reproductive coercion and women's health. *J Midwifery Womens Health*. 2018;63(5):518–525. CrossRef. Medline
20. Heise L, Ellsberg M, Gottmoeller M. A global overview of gender-based violence. *Int J Gynaecol Obstet*. 2002;78(Suppl 1):S5–S14. CrossRef. Medline

21. World Health Organization (WHO). *WHO Multi-Country Study on Women's Health and Domestic Violence against Women*. Geneva: WHO; 2005. https://www.who.int/reproductivehealth/publications/violence/24159358X/en/. Accessed September 24, 2019.
22. World Health Organization (WHO). *Global and Regional Estimates of Violence Against Women: Prevalence and Health Effects of Intimate Partner Violence and Non-Partner Sexual Violence*. Geneva: WHO; 2013. https://www.who.int/reproductivehealth/publications/violence/9789241564625/en/. Accessed September 24, 2019.
23. Black MC, Basile KC, Breiding MJ, et al. *The National Intimate Partner and Sexual Violence Survey (NISVS): 2010 Summary Report*. Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention; 2011. https://www.cdc.gov/violenceprevention/pdf/nisvs_report2010-a.pdf. Accessed September 24, 2019.
24. Clark LE, Allen RH, Goyal V, Raker C, Gottlieb AS. Reproductive coercion and co-occurring intimate partner violence in obstetrics and gynecology patients. *Am J Obstet Gynecol*. 2014;210(1):42.e1–e8. CrossRef. Medline
25. Hall M, Chappell LC, Parnell BL, Seed PT, Bewley S. Associations between intimate partner violence and termination of pregnancy: a systematic review and meta-analysis. *PLoS Med*. 2014;11(1):e1001581. CrossRef. Medline
26. Kazmerski T, McCauley HL, Jones K, et al. Use of reproductive and sexual health services among female family planning clinic clients exposed to partner violence and reproductive coercion. *Matern Child Health J*. 2015;19(7):1490–1496. CrossRef. Medline
27. Miller E, Silverman JG. Reproductive coercion and partner violence: implications for clinical assessment of unintended pregnancy. *Expert Rev Obstet Gynecol*. 2010;5(5):511–515. CrossRef. Medline
28. Miller E, Decker MR, McCauley HL, et al. Pregnancy coercion, intimate partner violence and unintended pregnancy. *Contraception*. 2010;81(4):316–322. CrossRef. Medline
29. Pearson E, Andersen KL, Biswas K, Chowdhury R, Sherman SG, Decker MR. Intimate partner violence and constraints to reproductive autonomy and reproductive health among women seeking abortion services in Bangladesh. *Int J Gynecol Obstet*. 2017;136(3):290–297. CrossRef. Medline
30. Jack CL, Mars M. Ethical considerations of mobile phone use by patients in KwaZulu-Natal: obstacles for mHealth? *Afr J Prim Heal Care Fam Med*. 2014;6(1):E1–E7. CrossRef. Medline
31. Freed D, Palmer J, Minchala D, Levy K, Ristenpart T, Dell N. Digital technologies and intimate partner violence: a qualitative analysis with multiple stakeholders. *Proc ACM Hum.-Comput. Interact*. 2017;1(1):1–22. CrossRef
32. Bonell C, Jamal F, Melendez-Torres GJ, Cummins S. "Dark logic": theorising the harmful consequences of public health interventions. *J Epidemiol Community Health*. 2015;69(1):95–98. CrossRef. Medline
33. Mayo-Wilson E, Fusco N, Li T, et al. Harms are assessed inconsistently and reported inadequately Part 1: Systematic adverse events. *J Clin Epidemiol*. 2019;113:11–19. CrossRef. Medline
34. Eckersberger E, Pearson E, Andersen K, et al. Developing mHealth messages to promote postmenstrual regulation contraceptive use in Bangladesh: participatory interview study. *JMIR mHealth UHealth*. 2017;5(12):e174. CrossRef. Medline
35. Jayaweera R, Baum S, Ramirez AM, Reyes Y, Chandrasekaran S, Gerdts C. Privacy by design, not default: centering user preferences in the development of sexual and reproductive health smartphone applications in five contexts. Paper presented at: International Conference on Family Planning; November 15, 2018; Kigali, Rwanda.
36. Smith C, Watkins J, Free C, Oreglia E. Development of an intervention to support reproductive health of women after seeking medical abortion in Cambodia. UK Research and Innovation website. https://gtr.ukri.org/projects?ref=AH%2FR006091%2F1. Accessed September 24, 2019.
37. McCarthy OL, Wazwaz O, Osorio Calderon V, et al. Development of an intervention delivered by mobile phone aimed at decreasing unintended pregnancy among young people in three lower middle income countries. *BMC Public Health*. 2018;18(1):576. CrossRef. Medline
38. McCarthy OL, Zghayyer H, Stavridis A, et al. A randomized controlled trial of an intervention delivered by mobile phone text message to increase the acceptability of effective contraception among young women in Palestine. *Trials*. 2019;20(1):228. CrossRef. Medline
39. Bailey J, McCarthy O, Carswell K, et al. The Sexunzipped website for sexual wellbeing for young people: early results of a pilot online RCT. *J Sex Med*. 2011;8:126.
40. Plourde K, L'Engle K. Design in equity. Paper presented at: International Conference on Family Planning. November 15, 2019; Kigali, Rwanda.
41. McCarthy OL, Osorio Calderon V, Makleff S, et al. An intervention delivered by app instant messaging to increase acceptability and use of effective contraception among young women in Bolivia: protocol of a randomized controlled trial. *JMIR Res Protoc*. 2017;6(12):252. CrossRef. Medline
42. McCarthy O, Ahamed I, Kulaeva F, et al. A randomized controlled trial of an intervention delivered by mobile phone app instant messaging to increase the acceptability of effective contraception among young people in Tajikistan. *Reprod Health*. 2018;15(1):28. CrossRef. Medline
43. Gonsalves L, Hindin MJ, Bayer A, et al. Protocol of an open, three-arm, individually randomized trial assessing the effect of delivering sexual and reproductive health information to young people (aged 13–24) in Kenya and Peru via mobile phones: Adolescent/youth reproductive mobile access and delivery initiative for love and life outcomes (ARMADILLO) study stage 2. *Reprod Health*. 2018;15(1):126. CrossRef. Medline
44. Lorenc T, Oliver K. Adverse effects of public health interventions: a conceptual framework. *J Epidemiol Community Health*. 2014;68(3):288–290. CrossRef. Medline
45. Jennings L, Gagliardi L. Influence of mHealth interventions on gender relations in developing countries: a systematic literature review. *Int J Equity Health*. 2013;12(1):85. CrossRef. Medline
46. Mulrenan C, Colombini M, Howard N, Kikuvi J, Mayhew SH. Exploring risk of experiencing intimate partner violence after HIV infection: a qualitative study among women with HIV attending postnatal services in Swaziland. *BMJ Open*. 2015;5(5):e006907. CrossRef. Medline
47. Kouyoumdjian FG, Findlay N, Schwandt M, Calzavara LM. A systematic review of the relationships between intimate partner violence and HIV/AIDS. *PLoS One*. 2013;8(11):e81044. CrossRef. Medline
48. Hill A, Pallitto C, McCleary-Sills J, Garcia-Moreno C. A systematic review and meta-analysis of intimate partner violence during pregnancy and selected birth outcomes. *Int J Gynaecol Obstet*. 2016;133(3):269–276. CrossRef. Medline
49. Stephenson R, Jadhav A, Winter A, Hindin M. Domestic violence and abortion among rural women in four Indian States. *Violence Against Women*. 2016;22(13):1642–1658. CrossRef. Medline
50. Ferreira A, Young T, Mathews C, Zunza M, Low N. Strategies for partner notification for sexually transmitted infections, including HIV. *Cochrane Database Syst Rev*. 2013;(10):CD002843. CrossRef. Medline
51. Mbuagbaw L, Thabane L, Ongolo-Zogo P, et al. The Cameroon Mobile Phone SMS (CAMPS) trial: a randomized trial of text

messaging versus usual care for adherence to antiretroviral therapy. *PLoS One.* 2012;7(12):e46909. CrossRef. Medline

52. Follingstad DR, Edmundson M. Is Psychological Abuse Reciprocal in Intimate Relationships? Data from a National Sample of American Adults. *J Fam Viol.* 2010;25(5):495–508. CrossRef
53. Follingstad DR, Bush HM. Measurement of intimate partner violence: a model for developing the gold standard. *Psychol Violence.* 2014; 4(4):369–383. CrossRef
54. Maiuro R. Sticks and stones may break my bones but names will also hurt me: Psychological abuse in domestically violent relationships. In: O'Leary K, Maiuro D, eds. *Psychological Abuse in Violent Domestic Relations*. New York, NY: Springer; 2001.
55. Hegarty K, Bush R, Sheehan M. The composite abuse scale: further development and assessment of reliability and validity of a multidimensional partner abuse measure in clinical meetings. *Violence Vict.* 2005;20(5):529–547. CrossRef. Medline
56. Jaffe AE, Steel AL, DiLillo D, Messman-Moore TL, Gratz KL. Characterizing sexual violence in intimate relationships: an examination of blame attributions and rape acknowledgment. *J Interpers Violence.* 2017:088626051772697. CrossRef. Medline
57. Dobash RP, Dobash RE, Wilson M, Daly M. The myth of sexual symmetry in marital violence. *Soc Probl.* 1992;39(1):71–91. CrossRef
58. Hamby S. The gender debate about intimate partner violence: solutions and dead ends. *Psychol Trauma Theory, Res Pract Policy.* 2009;1(1):24–34. CrossRef
59. Edwards PJ, Roberts I, Clarke MJ, et al. Methods to increase response to postal and electronic questionnaires. *Cochrane Database Syst Rev.* 2009;8(3):MR000008. CrossRef. Medline
60. Bacchus LJ, Buller AM, Ferrari G, Brzank P, Feder G. "It's always good to ask": a mixed methods study on the perceived role of sexual health practitioners asking gay and bisexual men about experiences of domestic violence and abuse. *J Mix Methods Res.* 2018;12(2): 221–243. CrossRef
61. Evans M, Gregory A, Feder G, Howarth E, Hegarty K. "Even 'Daily' is Not Enough": how well do we measure domestic violence and abuse?—a think-aloud study of a commonly used self-report scale. *Violence Vict.* 2016;31(1):3–26. CrossRef. Medline
62. Ellsberg M, Heise L. *Researching Violence Against Women: A Practical Guide for Researchers and Activists.* Washington, DC: World Health Organization, PATH; 2005.
63. World Health Organization (WHO). *Ethical and Safety Recommendations for Intervention Research on Violence against Women. Building on Lessons from the WHO Publication Putting Women First: Ethical and Safety Recommendations for Research on Domestic Violence against Women.* Geneva, Switzerland: WHO; 2016. https://www.who.int/reproductivehealth/publications/violence/intervention-research-vaw/en/.
64. Rahman A. Micro-credit initiatives for equitable and sustainable development: who pays? *World Dev.* 1999;27(1):67–82. CrossRef
65. García-Moreno C. Dilemmas and opportunities for an appropriate health-service response to violence against women. *Lancet.* 2002;359(9316):1509–1514. CrossRef. Medline
66. Colombini M, Mayhew S, Watts C. Health-sector responses to intimate partner violence in low- and middle-income settings: a review of current models, challenges and opportunities. *Bull World Health Organ.* 2008;86(8):635–642. CrossRef. Medline
67. Bacchus LJ, Bewley S, Vitolas CT, Aston G, Jordan P, Murray SF. Evaluation of a domestic violence intervention in the maternity and sexual health services of a UK hospital. *Reprod Health Matters.* 2010;18(36):147–157. CrossRef. Medline
68. Bacchus LJ, Bullock L, Sharps P, et al. Infusing technology into perinatal home visitation in the United States for women experiencing intimate partner violence: exploring the interpretive flexibility of an mHealth intervention. *J Med Internet Res.* 2016;18(11):e302. CrossRef. Medline
69. World Health Organization. *Responding to Intimate Partner Violence and Sexual Violence Against Women: WHO Clinical and Policy Guidelines.* Geneva, Switzerland: WHO; 2013. https://www.who.int/reproductivehealth/publications/violence/9789241548595/en/.
70. Bacchus LJ, Bullock L, Sharps P, et al. 'Opening the door': a qualitative interpretive study of women's experiences of being asked about intimate partner violence and receiving an intervention during perinatal home visits in rural and urban settings in the USA. *J Res Nurs.* 2016;21(5-6):345–364. CrossRef
71. Koziol-McLain J, Vandal AC, Wilson D, et al. Efficacy of a web-based safety decision aid for women experiencing intimate partner violence: randomized controlled trial. *J Med Internet Res.* 2018; 19(12):e426. CrossRef. Medline
72. Glass NE, Perrin NA, Hanson GC, et al. The longitudinal impact of an internet safety decision aid for abused women. *Am J Prev Med.* 2017;52(5):606–615. CrossRef. Medline
73. Jewkes R, Dartnall E. More research is needed on digital technologies in violence against women. *Lancet Public Heal.* 2019;4(6):e270–e271. CrossRef. Medline

Peer Reviewed

Received: June 28, 2019; **Accepted:** October 22, 2019

Cite this article as: Bacchus LJ, Reiss K, Church K, et al. Using digital technology for sexual and reproductive health: are programs adequately considering risk? *Glob Health Sci Pract.* 2019;7(4):507-514. https://doi.org/10.9745/GHSP-D-19-00239

GLOBAL HEALTH: SCIENCE AND PRACTICE
Dedicated to what works in global health programs
OPEN ACCESS

COMMENTARY

The Burden of Respiratory Syncytial Virus Infection in Adults and Reproductive-Aged Women

Bernard Gonik[a]

Currently available data on respiratory syncytial virus (RSV) disease burden in adults and reproductive-aged women are limited. These data are critically needed to assist in the advancement of strategies related to maternal RSV vaccination for the passive protection of their newborn children.

INTRODUCTION

In 2015, it was estimated that respiratory syncytial virus (RSV) infection resulted in 33.1 million episodes of acute lower respiratory infection worldwide, with 3.2 million hospitalizations and approximately 59,600 in-hospital related deaths.[1] These numbers likely underestimate the burden of disease because the overwhelming majority of cases are thought to occur in developing countries where surveillance data are even more limited.[1] Globally, RSV-associated acute lower respiratory infection is one of the leading causes of morbidity and mortality in children younger than 5 years.[1] Recently, this pathogen has also been recognized to cause significant disease in the elderly.[2] Despite 60 years of RSV research and vaccine exploration (Figure), there is only 1 approved intervention to prevent RSV infections. Palivizumab, a monoclonal antibody against the RSV fusion protein, is only indicated for preterm infants and children at high risk for RSV infections. No licensed vaccine currently exists. However, currently, 14 candidate vaccines are being tested in clinical trials.[4] Active vaccination of pregnant women in the third trimester is a particularly attractive approach because the most severe disease occurs within the first 6 months of life in their progeny. With current ongoing activities, approval of the first RSV vaccine for the prevention of RSV in all infants or perhaps the elderly is likely to occur in the next 6 years.

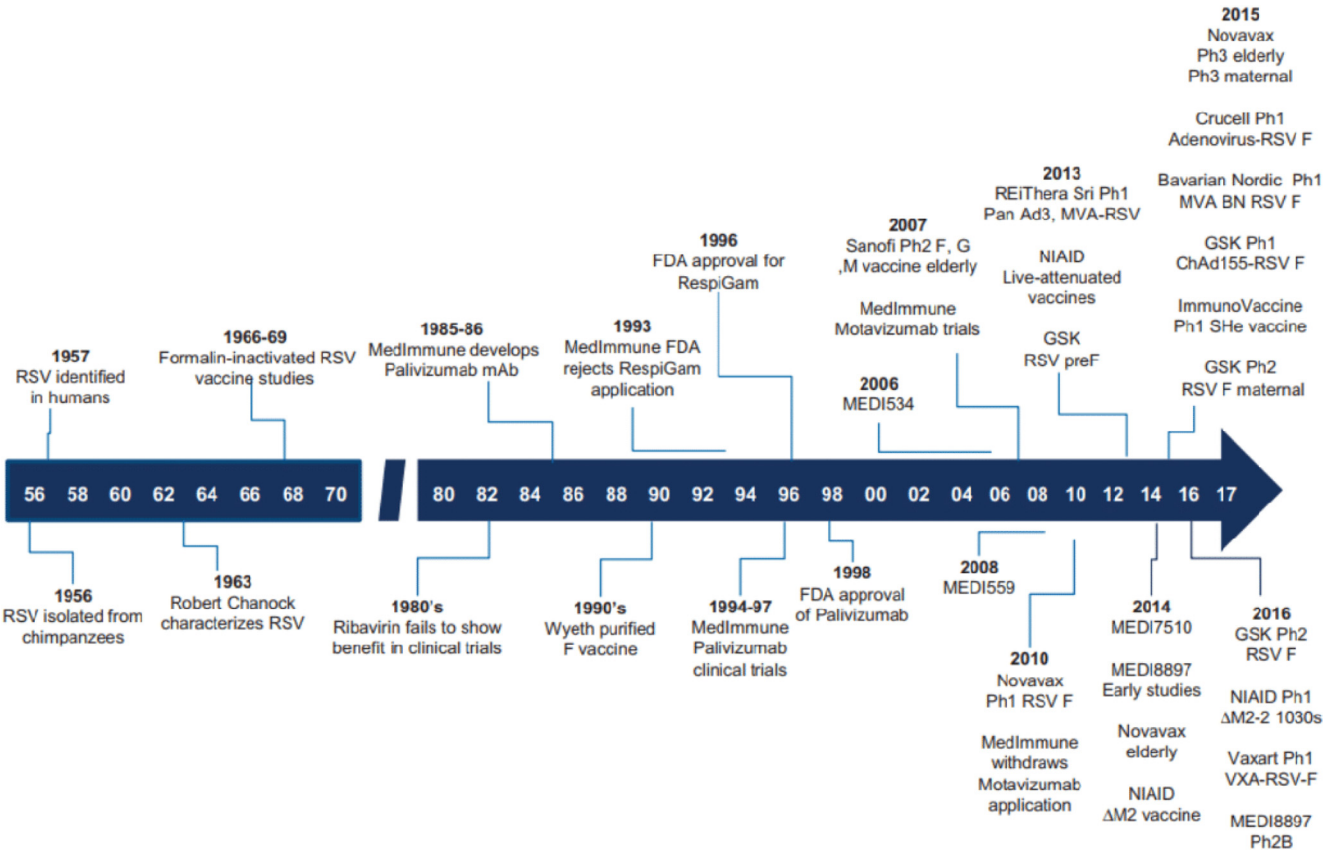
Why Characterize Disease Burden in a Non-Pediatric Population?

Although RSV vaccine discussion has focused on childhood infection, for several reasons, it may be important to better understand and characterize RSV disease burden in women of childbearing age. First, RSV disease is increasingly recognized as a significant contributor to adult respiratory illnesses worldwide.[5] Although pregnant women may receive an RSV vaccine to protect her child, other more direct benefits to the mother may be realized. Pregnancy is considered an immunologically attenuated state, and RSV infection during pregnancy is associated with more severe disease and adverse outcomes (e.g., fever, respiratory distress, preterm labor, hospitalization).[6] As a result, an RSV vaccine could improve maternal as well as infant outcomes and should be considered as part of cost/benefit analyses when planning vaccine introduction. Second, family members, including the mother, may be the source for neonate exposure to RSV; therefore, disease prevalence and modulation of that prevalence in an adult population is germane to this discussion. Third, recent data suggest the possibility of vertical RSV transmission during the prenatal period. When vertical transmission is induced in animal models, specific long-lasting alterations in immunologic and pulmonary functions have been demonstrated in the offspring of mothers acquiring the infection prenatally. Last, the degree of disease burden in adults may influence the timing of RSV vaccination, in that, if the burden is deemed low, the vaccine may be intentionally delayed to later in the third trimester to optimize antibody transfer and newborn protection.

To better understand and characterize RSV disease burden in adults, women of childbearing age, and pregnant women, we performed a literature search using English-language articles identified in the PubMed and Google Scholar databases, with additional resources identified by extraction from references within those documents. Keywords or phrases searched, used separately or in combination were: respiratory syncytial virus, RSV, epidemiology, adult respiratory disease, and pregnancy.

[a] Wayne State University School of Medicine, Detroit, MI, USA.
Correspondence to Bernard Gonik (bgonik@wayne.edu).

FIGURE. History of Respiratory Syncytial Virus Vaccine and Monoclonal Antibody Research[3]



Abbreviations: BN, Bavarian Nordic; ChAd, chimpanzee-derived adenovirus; F, respiratory syncytial virus fusion protein; FDA, United States Food and Drug Administration; GSK, Glaxo Smith Kline; mAb, monoclonal antibody; MEDI; MedImmune; MVA, modified vaccinia Ankara; NIAID, National Institute of Allergy and Infectious Diseases; PanAd, simian adenovirus; Ph, phase; RSV, respiratory syncytial virus; VXA, Vaxart.

RSV Biology

Human RSV is a member of the *Paramyxoviridae* family. It is a single stranded RNA virus that has 10 genes encoding 11 proteins. Two of these proteins, F and G, are particularly important for virus attachment and syncytia formation on target host cells. These 2 surface proteins are also the primary target antigens currently being investigated for both active and passive host immunologic protection. RSV is a virus that causes respiratory tract infections with a wide clinical spectrum. Outbreaks of RSV infections occur between fall and spring in temperate climates and tend to last up to 5 months. RSV isolates can be divided into 2 groups: group A and group B, based on antigenic and genetic characteristics. These 2 groups coexist in the human population, with group A being more prevalent. Most studies have not found significant clinical differences between both subtypes.[7]

Data for RSV in the adult population are limited by the paucity and quality of available studies.

RSV IN THE ADULT POPULATION

Data for RSV in the adult population are limited by the paucity and quality of available studies.[8–14] Regardless, certain general observations can be gleaned from the current literature and are summarized below. Data from an adult RSV-season surveillance program from 1975 to 1995 examined routinely obtained viral culture specimens and identified an RSV prevalence rate of 7% in the population.[8] The majority of subjects sampled were between 18 and 40 years old, and 65% were females. Of interest, 84% of culture-positive individuals were symptomatic, with 74% reporting upper respiratory symptoms, 26% with lower respiratory tract findings, and about half noting

subjective fever. Mean duration of viral shedding was 3.9 days, with a wide range (1–17 days). Compared to subjects with upper respiratory tract symptoms but who were RSV culture-negative, those with documented RSV infection had more prolonged symptoms and a greater frequency of concomitant lower respiratory findings. The investigators also noted that RSV infection was more commonly associated with acute exacerbations of chronic cardiopulmonary conditions, pulmonary function abnormalities, and increased mortality in immunocompromised hosts, although specific data were not presented. They also remarked on the propensity for RSV to spread among family members and those living in close quarters, although, again, data were not offered.

Internationally derived data show similar RSV infection rates in adults with respiratory symptoms. O'Shea et al. demonstrated an 11% prevalence of RSV in otherwise healthy male U.S. military recruits presenting with acute respiratory symptoms.[9] Using several isolation methodologies with different degrees of sensitivity for pathogen identification, these investigators also showed significant differences in detection rates, helping clarify the wide ranges reported in related studies. For example, Zambon et al. studied adults with upper respiratory tract infections in the United Kingdom and demonstrated a 20% RSV positivity rate using a sensitive polymerase chain reaction (PCR) technology.[10] In contrast, in a study in Kenya of patients with influenza-like illness or other acute respiratory symptoms, Bigogo et al. only detected a 4.2%–6.5% PCR-positive RSV occurrence.[11] However, this latter study was significantly compromised by an incomplete data collection process. These same investigators noted that adults with HIV infection were twice as likely as non-HIV patients to have RSV identified as a causative agent when presenting with acute respiratory illness.

One study attempted to diagnose RSV infection irrespective of symptoms. Munywoki et al. serially sampled 47 households (n=493) in rural Kenya throughout a complete RSV season regardless of symptomatology and found that 37% of the subjects had at least 1 RSV episode, with 58% symptomatic when RSV-positive by PCR.[12] Factors associated with symptomatic disease included younger age groups, higher viral loads, and extended shedding times. For individuals aged 15–40 years, 26% were identified to have at least 1 RSV episode, with only 24% being symptomatic. Fever was an uncommon symptom (14%), few sought medical care (16%), and none were hospitalized in the entire study cohort.

Complications of RSV Infection in Adults

The majority of RSV cases in the adult population are presumed to represent reinfection with this pathogen because primary acquisition in childhood is so common. Although most cases manifest clinically as mild to moderate upper respiratory disease, more severe complications have been recognized to occur. A prospective study in the United States of hospitalized adults with acute respiratory infections during the influenza season showed an RSV PCR prevalence rate of 8.7% in adults 50–64 years old.[13] Most of these individuals had at least 2 preexisting chronic comorbidities (e.g., congestive heart failure, chronic obstructive pulmonary disease, obesity) identified. In a multiinstitutional study in France that examined adults hospitalized with influenza-like illness, 4% were RSV-positive by multiplex reverse transcription-PCR.[14] Although the majority of these subjects were elderly, making it difficult to compare to a reproductive-age population, there was a 57% complication rate.[14] Pneumonia was identified in 44%, ICU admission was required for 15%, and the in-hospital case fatality rate was 8%. Luchsinger et al. studied 356 Chilean adults with community-acquired pneumonia and found an RSV prevalence of 13.4% by multiple detection methods including serology.[15] In 2 reviews of RSV disease and its complications in adults, experts noted that RSV accounts for between 2%–5% of adult pneumonia cases annually with an increase to 5%–15% of cases during the RSV season.[6,16] In the United States, it is estimated that between 11,000–17,000 adults die of RSV-related infection annually with 10-fold more admitted to the hospital with respiratory compromise. Han et al. reported the annual economic burden in the United States of adult inpatient care related to RSV approximated $150–680 million, with actual health care cost being much higher if non-pneumonia and outpatient cases were included in this calculation.[17] There are no similar cost estimates for other developed or low- to middle-income countries, but this will be an important consideration when planning for future interventions such as the introduction of RSV vaccines.

RSV in Reproductive-Aged Women and Pregnancy

Even fewer data on RSV are available in women of childbearing age or related to pregnancy.[18–24] In 2 different phase-2 clinical trials in healthy women 18–35 years old that looked at the potential for RSV vaccine use in reproductive-aged women,

Even fewer data are available on RSV in women of childbearing age or related to pregnancy.

investigators reported a 10%–15% Western blot RSV positivity rate at baseline, likely representing previous or recent exposures.[18,19] Interestingly, in both of these vaccine trials, which spanned 2 RSV transmission seasons, there was Western blot evidence of new RSV infection in 21% and 22%, respectively, in the placebo-arm study participants, unrelated to clinical symptomatology. In a study focused on influenza in rural Nepal, Chu et al. found a 2% RSV PCR positivity rate in pregnant or recently postpartum women with fever and acute respiratory symptoms.[20] Similarly, Chaw et al., in another prospective observational study of symptomatic pregnant women in Mongolia during the influenza season, noted a 2.4% RSV positive rate using a point-of-care immunoassay test.[21] Both of these latter 2 studies were limited by a primary focus on influenza-related symptoms, missed testing times, and other study design issues. In conjunction with a prospective maternal influenza vaccine trial, Madhi et al. retrospectively studied the incidence of RSV positivity by PCR in HIV-negative and HIV-positive South African gravidas and postpartum women.[22] Women presenting with respiratory illness from mid-pregnancy until 24 weeks postpartum underwent oral/nasal/pharyngeal swab sampling. The incidence of RSV associated illness (cases per 1000 person-months) during pregnancy in 2011 was 1.7 and 6.6 for HIV-negative and HIV-positive women, respectively. The incidence postpartum was identical for these 2 study populations at 2.3. The authors made additional observations, including the fact that RSV infection during pregnancy did not appear to be associated with any adverse pregnancy outcomes compared to the non-RSV infected groups. They expressed that their study likely underestimated the overall prevalence of RSV due to the focus on influenza-related disease, the variable incidence over RSV seasons (they noted the incidence in pregnancy in 2012 was 5.3 for HIV-negative pregnant women), and the fact that most RSV cases were identified during unsolicited medically attended visits which likely missed cases of milder disease. Hause et al. performed an outpatient, cross-sectional, surveillance study in pregnant women in their second and third trimesters in Houston, Texas. Of 81 subjects who presented with acute respiratory illness symptoms during the RSV season, 10% were identified to have RSV using tandem competitive-PCR.[23]

Regarding pregnancy complications associated with RSV infection, and despite the observations mentioned by Madhi et al.[22] in the study by Chu et al.,[20] of 7 patients with acute respiratory symptoms, fever, and RSV positivity, 2 (29%) experienced preterm birth. Given the very small numbers, statistical analysis comparing these subjects to the non-RSV infected study population (preterm birth rate 13%) was not possible. In a recent U.S. tertiary care center case report series, Wheeler et al. reported on 3 cases of symptomatic RSV infection during the antepartum period, with 2 of these patients requiring hospitalization and intubation for respiratory failure.[6] Both of these patients had other comorbidities including smoking and preexisting asthma.

In an international study examining RSV infection in pregnant women from high-income countries who were hospitalized, 48% were severe enough to require prolonged hospitalization and 38% were subsequently diagnosed to have pneumonia. The majority of these RSV cases were detected in the third trimester, and one-third had a preexisting health condition, most commonly asthma. Among women who did not deliver during that admission, there was a 29% preterm birth rate.[24] Limitations to this study, as is the case with most other studies on this topic, were the small number of patients sampled for RSV, a restriction to hospitalized patients only, and a study design that focused on influenza in pregnancy.

Although tangential in nature, data on influenza in pregnancy, another common viral pathogen, show that pregnant women are 7 times more likely to be hospitalized and 2 times more likely to die compared to non-pregnant women of reproductive age.[25] Lastly, and also tangential, pregnant women with pneumonia due to various pathogens, have an increased risk for several adverse pregnancy outcomes, including low birth weight, preterm birth and cesarean delivery.[26] As RSV continues to be investigated in this patient population, these types of adverse events need to be carefully monitored.

A novel concern is the concept of vertical transmission of RSV. Using a rodent model, investigators have recently reported evidence of RSV vertical transmission from mother to fetus, with persistence of infection in the offspring after birth.[27,28] These exposed and infected newborn pups showed evidence of prolonged altered viral immunity, dysregulation of neurotrophic pathways, and, most importantly, airway hyper-reactivity with RSV reinfection. Using human specimens, Fonceca and colleagues recently reported rescue of infectious RSV from cord blood mononuclear cells in 26 of 45 (57.7%) samples, with increases in recovery during winter months.[29] These findings await confirmation, but suggest another avenue of RSV pathogenesis that might be interrupted by preventing maternal disease via vaccination.

RELEVANCE TO POLICY MAKERS AND PROGRAM MANAGERS IN LOW- AND MIDDLE-INCOME COUNTRIES

Concomitant with ongoing basic and clinical science research exploring maternal RSV vaccination and passive newborn protection, policy makers and program managers need to begin examining the potential impact of these advances on their specific populations and prioritize the use of available resources to accomplish these goals. This effort is particularly germane in low- and middle-income communities where it is anticipated that disease burden is likely underestimated and the lack of existing infrastructures have compromised the collection of this critically needed information. Characterization of temporal and geographic patterns of RSV circulation by country and region will be needed, because this will directly impact the accuracy of disease burden estimates and influence planned interventional strategies. Using prospective active pathogen surveillance with sensitive diagnostic tools across a wide demographic range would be ideal, but these types of platforms will require partnerships with regional health authorities, international agencies, established research units, and global funders. The majority of RSV-related data pertaining to disease severity thus far has been collected in developed countries and in hospitalized or medically attended cases. These same data need to be extended to a broader patient base if accurate morbidity and mortality estimates are to be calculated for less developed countries. Additionally, more refined age group strata, with particular emphasis on extremes in age, will be needed to address the impact of vaccination in specific populations, such as pregnant women, very young children, and the elderly. Adequate data are lacking regarding the impact of RSV disease with specific comorbidities such as HIV disease. The role of environmental factors like smoke and indoor air pollution in RSV pathophysiology also need to be explored. More information on long-term disease outcomes, such as RSV reinfection and chronic wheezing, will help better define cost-effective and quality of life analyses that are critical components when considering the introduction of new therapeutic interventions. Last, if injectable vaccination is the preferred route of administration, a detailed analysis of population health care environments, drug delivery systems and pathways, and available personnel will need to be examined particularly in low-resource settings. A detailed description of this topic, including future research agenda items for consideration, can be found in a recently published document by PATH entitled "Advancing RSV Maternal Immunization: A Gap Analysis Report."[4]

CONCLUSIONS

In summary, this review highlights the limited and poor quality of the currently available surveillance and natural history RSV data in the adult population, particularly in low- to middle-income countries where the disease burden is likely to be significantly greater.[30] These data, specific to country or region, are critically needed to better understand the full extent of this pathogen's disease burden and to optimize future vaccine-related interventions. To this latter point, both the World Health Organization and the Bill & Melinda Gates Foundation have spearheaded research efforts to address these deficiencies. Despite significant methodologic concerns with the current data (such as varying approaches to pathogen identification, retrospective use of samples collected for other studies, etc.), it would be reasonable to speculate that the overall prevalence of symptomatic RSV infection in reproductive-aged or pregnant women is low, and severe morbidities are uncommon. Although isolated cases of severe lower respiratory disease in pregnancy have been reported, these instances may have been precipitated by other underlying comorbidities such as chronic lung disease and HIV infection. These potential relationships require further study. There is likely a much larger cadre of mildly symptomatic or asymptomatic pregnant women that remain of interest because of the risk for disease propagation and the potential for vertical transmission with long-term consequences in the progeny. As interventional investigations progress, the pregnant host should not be seen as simply a vehicle for passive antibody delivery to the fetus, but also as a potential beneficiary of active RSV vaccination.[31]

Policy makers and program managers should examine the potential impact of research advances in maternal RSV vaccination on their specific populations.

Acknowledgments: None.

Competing Interests: Data for this work was in part generated as a result of collaborative efforts with PATH, a global nonprofit health organization based in Seattle, Washington, for the development of a gap analysis for the Advancing Maternal Immunization project. Dr. Gonik is a site principal investigator for an ongoing Phase 3 randomized, placebo-controlled study to determine immunogenicity and safety of an RSV F nanoparticle vaccine in pregnant hosts by Novavax, Inc. (ClinicalTrials.gov identifier: NCT02624947).

REFERENCES

1. Shi T, McAllister DA, O'Brien KL, et al. Global, regional and national disease burden estimates of acute lower respiratory infections due to

respiratory syncytial virus in young children in 2015: a systematic review and modelling study. *Lancet.* 2017;390(10098):946–958. CrossRef. Medline

2. Lee N, Lui GC, Wong KT, et al. High morbidity and mortality in adults hospitalized for respiratory syncytial virus infections. *Clin Infect Dis.* 2013;57(8):1069–1077. CrossRef. Medline
3. Villafana T, Falloon J, Griffin MP, Zhu Q, Esser MT. Passive and active immunization against respiratory syncytial virus for the young and old. *Expert Rev Vaccines.* 2017; 16(7):737–746. CrossRef. Medline
4. PATH. *Advancing RSV Maternal Immunization: A Gap Analysis Report.* Seattle: PATH; 2018. https://path.azureedge.net/media/documents/Advancing_RSV_Maternal_Immunization__A_Gap_Analysis_Report.pdf. Accessed April 2, 2019.
5. Walsh EE, Falsey AR. Respiratory syncytial virus infection in adult populations. *Infect Disord Drug Targets.* 2012;12(2):98–102. Medline
6. Wheeler SM, Dotters-Katz S, Heine RP, Grotegut CA, Swamy GK. Maternal effects of respiratory syncytial virus infection during pregnancy. *Emerg Infect Dis.* 2015;21(11):1951–1955. CrossRef. Medline
7. Griffiths C, Drews SJ, Marchant DJ. Respiratory syncytial virus: infection, detection and new options for prevention and treatment. *Clin Microbiol Rev.* 2017;30(1):277–319. CrossRef. Medline
8. Hall CB, Long CE, Schnabel KC. Respiratory syncytial virus infections in previously healthy working adults. *Clin Infect Dis.* 2001; 33(6):792–796. CrossRef. Medline
9. O'Shea MK, Ryan MAK, Hawksworth AW, Alsip BJ, Gray GC. Symptomatic respiratory syncytial virus infection in previously healthy young adults living in a crowded military environment. *Clin Infect Dis.* 2005;41(3):311–317. CrossRef. Medline
10. Zambon MC, Stockton JD, Clewley JP, Fleming DM. Contribution of influenza and respiratory syncytial virus to community cases of influenza-like illness: an observational study. *Lancet.* 2001; 358(9291):1410–1416. CrossRef. Medline
11. Bigogo GM, Breiman RF, Feikin DR, et al. Epidemiology of respiratory syncytial virus infection in rural and urban Kenya. *J Infect Dis.* 2013;208 Suppl 3:S207–S216. CrossRef. Medline
12. Munywoki PK, Koech DC, Agoti CN, et al. Frequent asymptomatic respiratory syncytial virus infections during an epidemic in a rural Kenyan household cohort. *J Infect Dis.* 2015;212(11):1711–1718. CrossRef. Medline
13. Malosh RE, Martin ET, Callear AP, et al. Respiratory syncytial virus hospitalization in middle-aged and older adults. *J Clin Virol.* 2017;96:37–43. CrossRef. Medline
14. Loubet P, Lenzi N, Valette M, et al. Clinical characteristics and outcome of respiratory syncytial virus infection among adults hospitalized with influenza-like illness in France. *Clin Microbiol Infect.* 2017;23(4):253–259. CrossRef. Medline
15. Luchsinger V, Piedra PA, Ruiz M, et al. Role of neutralizing antibodies in adults with community-acquired pneumonia by respiratory syncytial virus. *Clin Infect Dis.* 2012;54(7):905–912. CrossRef. Medline
16. Falsey AR. Respiratory syncytial virus infection in adults. *Semin Respir Crit Care Med.* 2007;28(2):171–181. CrossRef. Medline
17. Han LL, Alexander JP, Anderson LJ. Respiratory syncytial virus pneumonia among the elderly: an assessment of disease burden. *J Infect Dis.* 1999;179(1):25–30. CrossRef. Medline
18. August A, Glenn GM, Kpamegan E, et al. A phase 2 randomized, observer-blind, placebo-controlled, dose-ranging trial of aluminum-adjuvanted respiratory syncytial virus F particle vaccine formulations in healthy women of childbearing age. *Vaccine.* 2017;35 (30):3749–3759. CrossRef. Medline
19. Glenn GM, Fries LF, Thomas DN, et al. A randomized, blinded, controlled, dose-ranging study of a respiratory syncytial virus recombinant fusion (F) nanoparticle vaccine in healthy women of childbearing age. *J Infect Dis.* 2016;213(3):411–422. CrossRef. Medline
20. Chu HY, Katz J, Tielsch J, et al. Clinical presentation and birth outcomes associated with respiratory syncytial virus infection in pregnancy. *PLoS One.* 2016;11(3):e0152015. CrossRef. Medline
21. Chaw L, Kamigaki T, Burmaa A, et al. Burden of influenza and respiratory syncytial virus infection in pregnant women and infants under 6 months in Mongolia: A prospective cohort study. Cowling BJ, ed. *PLoS One.* 2016;11(2):e0148421. CrossRef. Medline
22. Madhi SA, Cutland CL, Downs S, et al. Burden of respiratory syncytial virus infection in South African human immunodeficiency virus (HIV)-infected and HIV-uninfected pregnant and postpartum women: A longitudinal cohort study. *Clin Infect Dis.* 2018;66(11):1658–1665. CrossRef. Medline
23. Hause AM, Avadhanula V, Maccato ML, et al. A cross-sectional surveillance study of the frequent and etiology of acute respiratory illness among pregnant women. *J Infect Dis.* 2018;218(4):528–535. CrossRef. Medline
24. Regan AK, Klein NP, Langley G, et al. Respiratory syncytial virus hospitalization during pregnancy in 4 high-income countries, 2010–2016. *Clin Infect Dis.* 2018;67(12):1915–1918. CrossRef. Medline
25. Van Kerkhove MD, Vandemaele KA, Shinde V, et al. Risk factors for severe outcomes following 2009 influenza A (H1N1) infection: a global pooled analysis. *PLoS Med.* 2011;8(7):e1001053. CrossRef. Medline
26. Chen YH, Keller J, Wang IT, Lin CC, Lin HC. Pneumonia and pregnancy outcomes: a nationwide population-based study. *Am J Obstet Gynecol.* 2012;207(4):288.e1–288.e7. CrossRef. Medline
27. Piedimonte G, Walton C, Samsell L. Vertical transmission of respiratory syncytial virus modulates pre- and postnatal innervation and reactivity of rat airways. *PLoS One.* 2013;8(4):e61309. CrossRef. Medline
28. Brown PM, Harford TJ, Agrawal V, Yen-Lieberman B, Rezaee F, Piedimonte G. Prenatal exposure to respiratory syncytial virus alters postnatal immunity and airway smooth muscle contractility during early-life reinfections. *PLoS One.* 2017;12(2):e0168786. CrossRef. Medline
29. Fonceca AM, Chopra A, Levy A, et al. Infective respiratory syncytial virus is present in human cord blood samples and most prevalent during winter months. *PLoS One.* 2017;12(4):e0173738. CrossRef. Medline
30. Poletti P, Merler S, Ajelli M, et al. Evaluating vaccination strategies for reducing infant respiratory syncytial virus infection in low-income settings. *BMC Med.* 2015;13:49. CrossRef. Medline
31. Chamberlain AT, Lavery JV, White A, Omer SB. Ethics of maternal vaccination. *Science.* 2017;358(6362):452–453. CrossRef. Medline

Peer Reviewed

Received: April 4, 2019; **Accepted:** October 1, 2019; **First Published Online:** November 27, 2019

Cite this article as: Gonik B. The burden of respiratory syncytial virus infection in adults and reproductive-aged women. *Glob Health Sci Pract.* 2019; 7(4):515-520. https://doi.org/10.9745/GHSP-D-19-00121

GLOBAL HEALTH: SCIENCE AND PRACTICE
Dedicated to what works in global health programs
OPEN ACCESS

ORIGINAL ARTICLE

Effectiveness of an Electronic Partogram: A Mixed-Method, Quasi-Experimental Study Among Skilled Birth Attendants in Kenya

Harshadkumar Sanghvi,[a] Diwakar Mohan,[b] Lindsay Litwin,[a] Eva Bazant,[a] Patricia Gomez,[a] Tara MacDowell,[a] Levis Onsase,[a] Valentino Wabwile,[a] Charles Waka,[a] Zahida Qureshi,[c] Eunice Omanga,[a] Anthony Gichangi,[a] Ruth Muia[d]

Use of the electronic partogram, a digital labor-support application, is associated with improved fetal outcomes and greater use of interventions to maintain normal labor compared to the paper partograph.

ABSTRACT

Background: Timely identification and management of intrapartum complications could significantly reduce maternal deaths, intrapartum stillbirths, and newborn deaths due to hypoxia. The World Health Organization (WHO) identifies monitoring of labor using the paper partograph as a high-priority intervention for identifying abnormities in labor and fetal well-being. This article describes a mixed-method, quasi-experimental study to assess the effectiveness of an Android tablet-based electronic, labor clinical decision-support application (ePartogram) in limited-resource settings.

Methods: The study, conducted in Kenya from October 2016 to May 2017, allocated 12 hospitals and health centers to an intervention (ePartogram) or comparison (paper partograph) group. Skilled birth attendants (SBAs) in both groups received a 2-day refresher training in labor management and partograph use. The intervention group received an additional 1-day orientation on use and care of the Android-based ePartogram app. All outcomes except one compare post-ePartogram intervention versus paper partograph controls. The exception is outcome of early perinatal mortality pre- and post-ePartogram introduction in intervention sites compared to control sites. We used log binomial regression to analyze the primary outcome of the study, suboptimal fetal outcomes. We also analyzed for secondary outcomes (SBAs performing recommended actions), and conducted in-depth interviews with facility in-charges and SBAs to ascertain acceptability and adoptability of the ePartogram.

Results: We compared data from 842 clients in active labor using ePartograms with data from 1,042 clients monitored using a paper partograph. SBAs using ePartograms were more likely than those using paper partographs to take action to maintain normal labor, such as ambulation, feeding, and fluid intake, and to address abnormal measurements of fetal well-being (14.7% versus 5.3%, adjusted relative risk=4.00, 95% confidence interval [CI]=1.95–8.19). Use of the ePartogram was associated with a 56% (95% CI=27%–73%) lower likelihood of suboptimal fetal outcomes than the paper partograph. Users of the ePartogram were more likely to be compliant with routine labor observations. SBAs stated that the technology was easy to use but raised concerns about its use at high-volume sites. Further research is needed to evaluate costs and benefit and to incorporate recent WHO guidance on labor management.

Conclusion: ePartogram use was associated with improvements in adherence to recommendations for routine labor care and a reduction in adverse fetal outcomes, with providers reporting adoptability without undue effort. Continued development of the ePartogram, including incorporating new clinical rules from the 2018 WHO recommendations on intrapartum care, will improve labor monitoring and quality care at all health system levels.

INTRODUCTION

In 2015, the World Health Organization (WHO) estimated that 303,000 maternal deaths occur worldwide each year.[1] A 2013 global burden of disease study estimated that 6.4% of maternal deaths annually were due to obstructed labor.[2] In addition, 1.3 million intrapartum stillbirths and 904,000 newborn deaths due to hypoxia occur each year.[3,4] Timely identification and management of intrapartum complications could prevent many of these deaths.[5] With the global impetus toward universal health coverage, more women are choosing to give birth in health facilities; however, health outcomes will not improve unless service quality is assured.[6] WHO identifies monitoring of labor to guide timely, appropriate

[a] Jhpiego, Baltimore, MD, USA.
[b] Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA.
[c] University of Nairobi, Nairobi, Kenya.
[d] Kenya Ministry of Health, Nairobi, Kenya.
Correspondence to Patricia Gomez (patricia.gomez@jhpiego.org).

actions as a high-priority quality improvement intervention.[7]

Labor management requires skilled birth attendants (SBAs) to record periodic observations of maternal and fetal well-being, use these data to distinguish normal from abnormal progress, and predict and plan next steps over the course of labor. Interpreting a single measurement such as fetal heart rate is relatively simple, but evaluating combinations of measurements (e.g., labor progression in relation to frequency and duration of contractions) is complex. In optimal labor management, women progressing normally are supported by ambulation, oral fluids, feeding, and presence of a companion of choice, and unwarranted use or overuse of interventions, such as artificial rupture of membranes and augmentation, are avoided. When labor abnormalities occur, appropriate actions are taken.

The paper partograph is the most commonly available labor-monitoring tool, used by health professionals and recommended by WHO for active labor.[8] The WHO partograph is a graphic representation of measures of fetal and maternal well-being and labor progression that facilitates identification of obstetric and fetal complications. Routine use of the paper partograph in low- and middle-income countries is inconsistent,[9,10] and in many settings, SBAs complete partographs retrospectively for recordkeeping purposes only.[11–14]

A study in Kenya in 2010 found that "the partograph was widely used during labor (in 88% of 442 cases) and was initiated at the correct time in more than 90% of the cases; however, all components were completed in only 58% of cases."[15] A 2018 Cochrane review concluded that the quality of existing evidence was too low to determine whether using a paper partograph compared to nonuse of a paper partograph affected the rate of cesarean delivery or incidence of low Apgar scores.[16] Using a partograph may make little difference in labor length (low-quality evidence) or the number of women who receive oxytocin to accelerate labor (moderate-quality evidence).[16] Bedwell et al. concluded from a realist review that although the paper partograph appears to be accepted, evidence suggests that it is not being used in practice as anticipated and thus is not reaching its potential in improving outcomes.[17]

Several developers have focused on low-cost digital applications to address deficiencies in the paper partograph, improve recordkeeping, support decision making, and enhance quality of care during labor and delivery.[18–21] In one of the first published evaluations of digital labor-support applications, Litwin et al. reported on the use of an Android tablet-electronic partogram application (ePartogram) in Zanzibar.[22] Health workers quickly became competent and confident in using the ePartogram application on a tablet and believed its use improved timeliness of care and supported decision making.

This study seeks to assess the effectiveness of ePartogram use on health outcomes in limited-resource settings and to ascertain the acceptability and adoptability of the ePartogram based on health workers' experiences.

METHODS

Study Design

This was a mixed-method, quasi-experimental evaluation of labor management outcomes. Our intervention included a 2-day refresher training for SBAs and supervisors in control and intervention sites and introduction and training on use of the ePartogram in intervention sites. All outcomes except one were during the intervention period, comparing the ePartogram intervention. The exception was outcome of early perinatal mortality before and during ePartogram introduction in intervention sites compared to control sites.

Study Setting and Sites

The study was conducted from October 2016 to May 2017 in 12 health facilities serving 2 counties, Kisumu in western Kenya and Meru in eastern Kenya. The 2 tertiary care facilities, 1 per county, were allocated to receive the intervention. The other facilities were selected after matching on characteristics like birth volume, staffing level, facility type (public or private), and provision of basic and/or comprehensive emergency obstetric and newborn care (BEmONC and CEmONC, respectively). We allocated 2 large CEmONC facilities to the ePartogram group. The remaining 10 facilities had similar overall delivery rates and were randomly allocated to the intervention or comparison group. Thus, the ePartogram group had 2 large CEmONC and 4 BEmONC facilities, and the paper partograph group had 4 small CEmONC and 2 BEmONC facilities. The Kenyan public health care structure has large central tertiary hospitals, usually 1 per county, and smaller district hospitals and health centers. The larger hospitals usually have a specialist obstetrician-gynecologist and many midwives, and health centers often only have 1 or 2 midwives. Often care is given by providers who do not fit the WHO

definition of SBAs. Public services are complemented by faith-based institutions and private maternities.

Kenyan health facilities use the WHO modified partograph with alert and action lines; a partograph is started once the woman's cervix is 4 cm dilated. SBAs are supposed to fill out partographs, but many Kenyan health facilities are poorly staffed, and sometimes partographs are completed by nurses, students, and other non-SBAS who are deployed in labor wards.

The Intervention

The ePartogram is an Android tablet-based application developed using human-centered design between 2011 and 2017 to address many challenges of monitoring labor with the paper partograph. The ePartogram is a clinical decision-support tool with algorithms and clinical rules that are based on WHO guidance for managing normal and complicated labor. The app provides auditory reminders to prompt providers to take measurements when due and provides visual and auditory alerts when clinical rules are triggered that predict complications (via low-level alerts) or detect complications (via high-level alerts) (Figure 1). The app allows for increased data entry efficiency by automatically graphing data and storing all client files within the application. In addition, retrospective entry of data is prevented. The decision-supporting software is based on 77 clinical rules based on measures of fetal and maternal well-being, progress of labor, and expected trends as labor progresses.

The ePartogram app can handle multiple patients who a provider may be taking care of. A dashboard displays all the patients in a provider's care and highlights those that need immediate attention. In intervention sites, supervisors had access to digitally transmitted ePartogram data anytime on a tablet device if they chose to review and act on them. In control sites, supervisors had access to paper partographs when they visited or were called to the labor ward, which was the supervision norm. Remote supervision off-site with the ePartogram was possible, but we did not activate this function during the study. Data were stored securely in the cloud and were available in real time. The app also prompted the provider to record actions and interventions they took. The app can be tailored to the needs of the country by adding or removing clinical rules. The technical details of the ePartogram are described elsewhere.[22]

The usability and functionality of the ePartogram underwent field validation exercises in Kenya and Tanzania and a feasibility study in Zanzibar.[22] Qualitative feedback from the Zanzibar study informed additional functional enhancements to the ePartogram, including the ability to print a paper record of the ePartogram, inclusion of a dashboard for supervisors and managers, the ability to digitally transfer the full labor record to referral sites, and use of more intuitive icons on the app.

Participant Recruitment, Consent, and Training

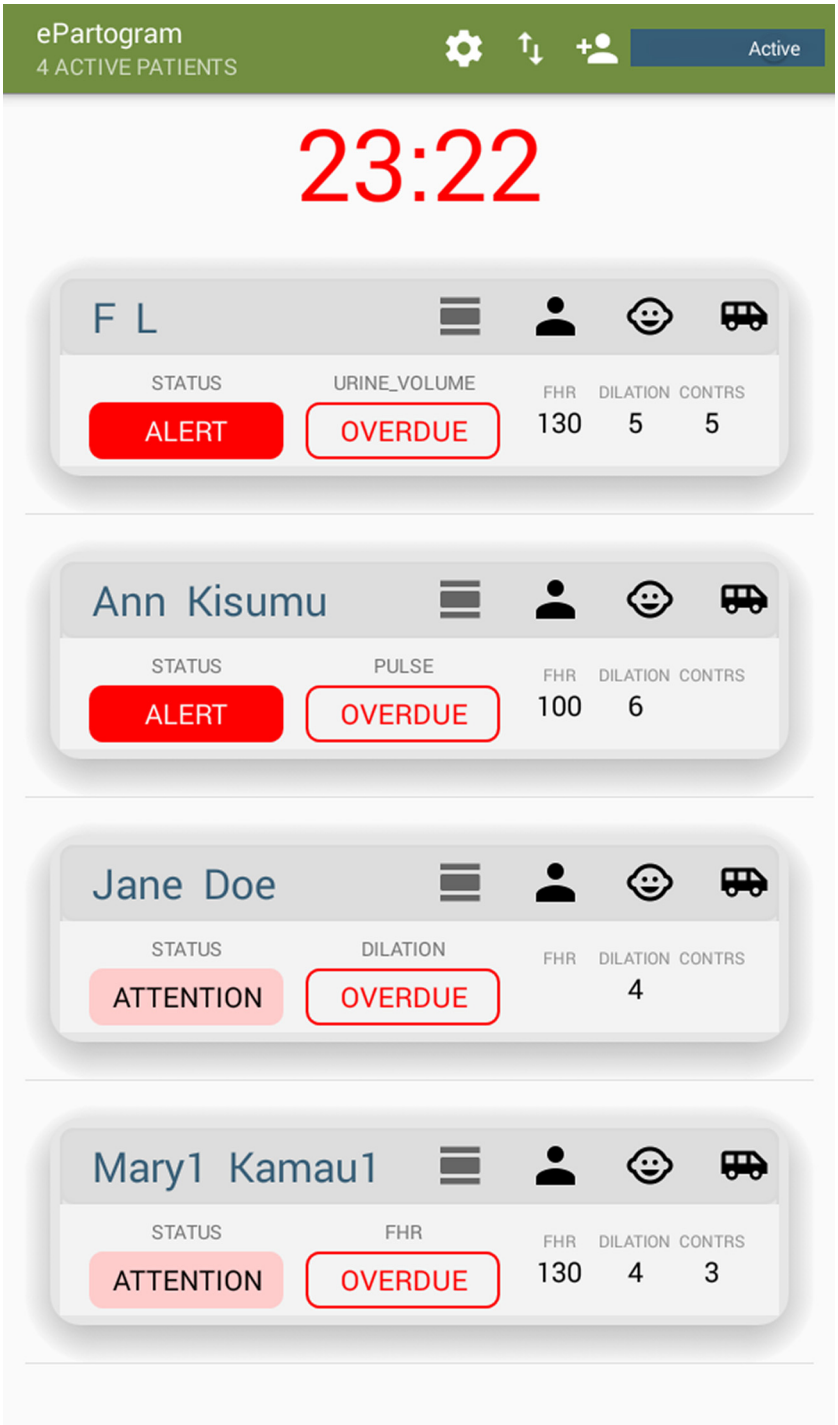
The ePartogram app was developed to address many challenges of monitoring labor with the paper partograph and improve decision making.

Study coinvestigators trained 9 master trainers during a 2-day workshop. This training-of-trainers workshop included interactive presentations and case studies, and it facilitated practice using both the paper partograph and ePartogram. Clinicians meeting the WHO definition of SBA (that is, a doctor, nurse, or midwife "trained to proficiency in the skills needed to manage normal [uncomplicated] pregnancies, childbirth and the immediate postnatal period, and in the identification, management and referral of complications in women and newborns"[23]) and caring for women in active labor at study sites were recruited. We did not include midwifery students or any unqualified birth attendant even though they often provide labor care in these facilities. All SBAs in a selected facility were recruited to the same study group and gave individual written consent to participate.

All participating skilled birth attendants and supervisors completed a 2-day labor management update.

All participating SBAs and supervisors completed a 2-day labor management update conducted by master trainers and coinvestigators. This training included case studies on using the paper partograph, decision making, and managing normal labor and common labor complications according to WHO and Kenya Ministry of Health guidelines. Additional topics included respectful maternity care; recognition and management of fetal, maternal, and labor progression abnormalities; fetal distress; pre-eclampsia/eclampsia; and fever. Trainees were also refreshed on when supervisors should be contacted in response to labor abnormalities identified, the standard operating procedures for handing over patients as shifts changed, and how to fill out facility birth and outcome registers. We had an additional layer of scrutiny for the ePartogram group only, although this did not involve helping with clinical issues and dealt only with fixing the occasional technology-related issues, such as failure of supervisor tablets to sync with SBA tablets.

FIGURE 1. ePartogram Dashboard Screenshot

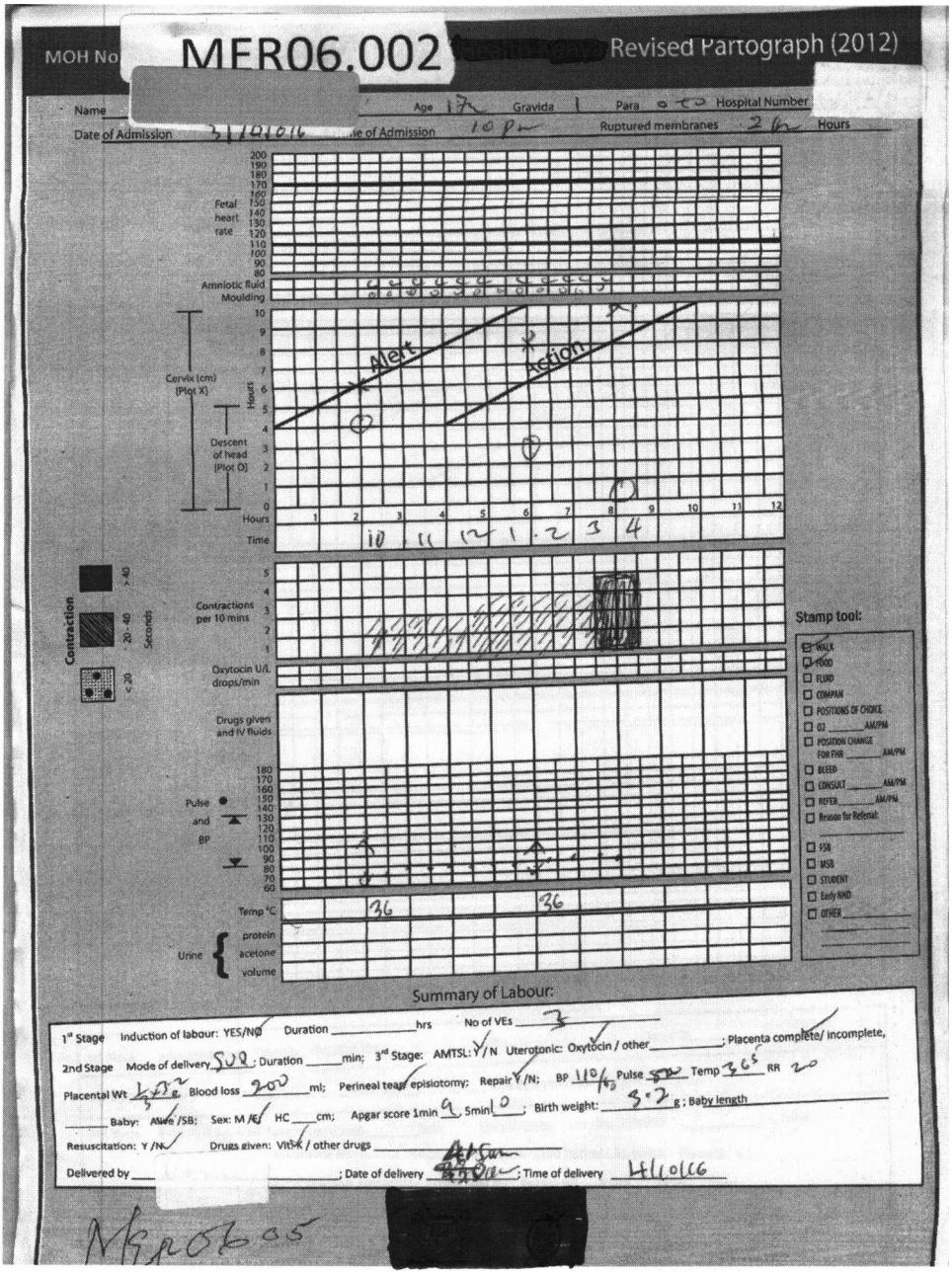


Dashboard shows patients under care of a provider, alerts, and overdue assessments.

We used guided case studies to instruct SBAs on how to interpret clinical information displayed on the partograph. The content and approach for the 2-day labor management training was identical for both study arms. The ePartogram provided prompts to record interventions and actions that the SBA takes, and because this was not available on the paper partograph, we included an action-taken sticker on all paper partographs so that actions could be recorded (Figure 2).

Intervention group participants were also trained in the use of the ePartogram application, maintenance of the tablets, and handover protocols. During the day-long ePartogram training, we

FIGURE 2. Completed Paper Partograph With Stamp Tool to Record Interventions



provided step-by-step didactic training on the application. We also trained the participants on the standard operating procedures for ePartogram use, storage, and cleaning and on how to print from the ePartogram.

We excluded SBAs who did not pass the written assessment during the update training, did not pass the written or skills-based assessments, or did not complete the training. Overall, 72 SBAs were trained in the intervention group, of whom 69 passed the assessments and consented to participate the study. Of the 42 SBAs trained in the comparison group, 1 withdrew from the study. At all study sites, women in active labor received care according to global and national standards. SBAs in intervention facilities used the ePartogram to document measurements; back-up paper partographs were available at each site. In the comparison facilities, SBAs used paper partographs to monitor women in labor. Client names and identifiers from the ePartograms and paper partographs were removed before partographs were scanned. We only used the partographs and labor records filled out by trained, consenting providers for this study.

Outcomes

Study outcomes relate to compliance with globally recommended labor-monitoring practices and recording the measurements obtained on the paper partograph and ePartogram; actions taken to maintain normal labor; detection of and actions taken to address deviations from normal labor and complications; and client outcomes. Table 1 describes the indicators used as outcome measures.

Study Sample Size

When the study was originally conceived, the sample size was calculated at 2,600 based on detecting differences in outcomes among laboring women whose parameters placed them to the right of the alert line (denoting abnormal progress). As the action line has in recent times been increasingly called into question, we abandoned a sample size based on that parameter. In addition, recruitment was disrupted by a health worker strike, resulting in a significantly larger proportion of patients arriving late in labor. Hence, we have presented a power analysis of the final effective sample size.

Overall, 1,884 paper partographs and ePartograms were collected. Analysis was done on a total sample of 1,884 and on subsets of this sample where some variables were missing. For some outcomes, we restricted the study analysis to those with 2 or more entries, giving a sample of 1,609. For fetal outcomes, after accounting for missing data, we analyzed a final sample size of 1,405. We performed a post hoc calculation using a type I error of 0.05 and power of 80% to assess the detectable difference. The primary outcome was a composite fetal

TABLE 1. Outcome Measures Used in the Mixed-Method Quasi-Experimental Study of the Effectiveness an Electronic Partogram

Primary Outcome Measure
1. Percentage of ePartogram/partographs showing fetus/newborn with a suboptimal fetal/newborn outcome (defined by presence of fresh stillbirth; newborn Apgar score of 5 or below at 1 minute, or 7 or below at 5 minutes; or newborn resuscitation needed) as recorded on the ePartogram/partograph by the SBA.
Secondary Outcome Measures
1. Percentage of ePartogram/partographs with a suboptimal maternal outcome (defined by presence of retained placenta, blood loss greater than or equal to 500 ml, systolic blood pressure less than 90 mm Hg or equal to/greater than 140 mm Hg, diastolic blood pressure less than 60 mm Hg or equal to/greater than 90 mm Hg, and pulse less than 60 beats per minute or equal to/greater than 100 beats per minute) as recorded by the SBA on the ePartogram/partograph.
2. Percentage of ePartogram/partographs with any action recorded on the ePartogram/partograph to maintain normal labor, among all partographs or ePartograms. Actions to maintain normal labor include ambulation and encouragement to eat or drink.
3. Percentage of ePartogram/partographs with any action recorded on the ePartogram/partograph to address any sign of deviation from normal labor, among all paper partographs or ePartograms to support detection, decision making, and action to address deviations from normal labor and complications arising during labor. Actions included providing oxygen, changing the position of the laboring woman in response to fetal heart rate abnormalities, checking for bleeding, consulting with a supervisor, referring a client to another facility, and augmenting labor during the first stage.
4. Incidence of fresh stillbirth and neonatal death within 24 hours among all births in a month, according to aggregate monthly routine facility data recorded on facility registers by SBAs.

Abbreviations: ePartogram, electronic partogram; SBA, skilled birth attendant.

outcome with a control group prevalence of 18% and an intraclass correlation coefficient of 0.11. The sample size would have been sufficient to detect a 16% difference between the intervention and comparison groups.

Quantitative Data Collection and Extraction

A trained health records information officer at each facility de-identified and scanned each completed partograph into the study database for cleaning and analysis. All paper partographs had an additional data collection tool stamped on them (Figure 2) or integrated into the electronic app (ePartogram), which asked whether the SBA instructed the client to walk/ambulate, eat food, or drink water during labor; gave the client fluid; encouraged the presence of a birth companion; administered oxygen; changed the client's position; checked for bleeding; consulted a specialist; referred the client to another site; or performed any other clinical action. Data were also collected about the number of births and maternal and newborn outcomes reported monthly in the maternity ward register. Once research assistants and/or study staff members received these files, they checked them to ensure they were legible, scanned correctly, and titled correctly and then saved the files in a master folder. Research assistants reviewed these scanned partographs and extracted data into the clinical and translational research software, Research Electronic Data Capture (REDCap).[24] Research assistants were not blinded to use of the ePartogram/partograph. Each research assistant's data extraction was validated by a sample of charts that were double entered by another assistant as part of the extraction training process. All data were abstracted by the researchers using REDCap.

Analysis of Individual Partographs

Data were analyzed in Stata version 14.2.[25] For this study, we assumed that missing data in paper partograph and ePartogram measurements were not missing at random; hence, we did not impute them. All estimates presented were based on complete case analysis. We used frequencies and percentages for categorical variables and median to present descriptive statistics for client and facility characteristics across the study groups. We tested hypotheses using nonparametric methods with Fisher exact tests for categorical variables and the test of medians for continuous variables. To assess the intervention's overall impact, we created a log binomial model with fetal outcome as the dependent variable and ePartogram use as the independent (intervention) variable. We used facility and individual level variables as control variables as part of multivariable regression model to account for the baseline differences between the intervention and comparison groups. The control variables included the type of facility (CEmONC versus BEmONC), affiliation (public versus faith-based), and number of providers at the facility. We also controlled for a nationwide health worker strike that disrupted health services at study sites between December 2016 and February 2017. At the individual level, we used parity (primipara versus multipara) and cervical dilation at admission ($\leq$5 cm versus >5 cm). The adjusted relative risks are presented as the effect size of the ePartogram intervention along with their 95% confidence intervals (CIs). Because laboring women may have been attended to by multiple providers from the facility, we accounted for clustering only at the level of the facilities using Huber-White sandwich estimators.

Analysis of Facility Registers

To estimate the intervention's effect, we compared data obtained from facility registers from 6 months before the study (May–October 2016) and after the implementation of the study (May 2017). All SBAs and supervisors were updated on how to correctly fill out facility registers during the training. In addition, register data were extracted and collected weekly from study-trained health records officers. No other effort was made to ensure completeness or validity of these data. The data points included the monthly aggregates for total births, number of stillbirths, and number of newborn deaths in the first 24 hours for each facility. The final analysis included register data from the 12 facilities on births over 12 months (2 facilities had 0 births in a month). The outcome variable was early perinatal mortality, defined as early neonatal (<24 hours) death and fresh stillbirth over a denominator of the sum of live births and fresh stillbirths. We analyzed for difference between intervention and comparison groups before and after the study by fitting a population-averaged generalized linear model using generalized estimating equations with a Poisson distribution, log link, and exchangeable correlation structure. The models were adjusted using facility-level clustering and adjusts for facility variables of affiliation (public versus private), capacity to address emergencies (CEmONC versus BEmONC), presence of a medical or clinical officer (yes/no), percentage of providers trained in labor management by study staff (≥75% or <75%),

median years of provider experience, and whether a health worker strike was occurring in the month in public facilities (yes/no). Incident rate ratios (IRRs) with 95% CIs for before and after rates in each study group are presented along with the difference-in-difference estimator (ratio of IRRs).

Qualitative Data Collection and Analysis

Four facility in-charges and 28 SBAs from both arms participated in in-depth interviews lasting 30 to 35 minutes. Interviews were conducted in English, recorded, and transcribed. Data collectors were implementing agency staff trained in qualitative research data collection procedures using a data collector training guide[26] and study ethics. We coded the qualitative data in Atlas.ti version 7 software,[27] and created a codebook based on field guide topics and themes that emerged from the interviews. Interview data saturation was deemed to have been reached, as no new themes emerged in the final interview and all themes were mentioned to some extent in all interviews. The analysis followed the framework analysis process recommended by Ritchie et al.[28] We identified themes and subthemes to align with research questions; we describe these later using illustrative quotations.

Ethical and Safety Considerations

The study was approved by the Johns Hopkins Bloomberg School of Public Health Institutional Review Board (JHSPH IRB #6958) and the Kenya Medical Research Institute (KEMRI protocol #530). SBAs gave informed written consent before participation. Women cared for using the ePartogram did not have a paper partograph plotted, but on completion of labor, a printout was made for the record and for de-identification before scanning to the study database. The JHSPH and KEMRI Institutional Review Boards required that providers participating in the study give consent but did not require that women cared for during labor provide consent.

RESULTS

We compared data from 842 clients in active labor in the ePartogram group with 1,042 clients in the paper partograph group.

Table 2 compares the facility characteristics of ePartogram and paper partograph sites, and Table 3 compares the client characteristics. Among facility characteristics, there were nonsignificant differences between paper partograph and ePartogram groups in median number of providers available, volume of deliveries, and proportion of all women that were recruited in the study. Although we used partographs only from study-trained providers, the ePartogram facilities had 94.5% of SBAs trained for the study compared to 65% of SBAs in paper partograph facilities, a significant difference (P=.02).

The ePartogram group consisted of public facilities only, with no faith-based facilities, and also had the 2 largest public facilities; 35.6% of recruitment was in BEmONC facilities to 30.4% in the paper partograph group. The study was disrupted by a prolonged health worker strike, during which 53.9% of the recruitment in the ePartogram group compared to 34.1% in the paper partograph group occurred, and this difference was significant. For cervical dilation at admission, 74.1% of ePartogram labors versus 63.4% of paper partograph labors were already beyond 5 cm dilated on admission, which may also partly, but not completely, explain the difference in median duration of the first stage of labor recorded at 120 minutes for the ePartogram group versus 555 minutes for the paper partograph group.

Table 4 compares the proportion of labors in which clinical rules showing abnormalities were triggered. In the paper partograph group, the SBA should recognize an abnormality in clinical rules during ongoing use of the partograph; in the ePartogram group, an abnormality in clinical signs will trigger a visual and an audible alert reminding SBAs to perform measurements when due or to take clinical actions. Clinical rules for fetal well-being (fetal heart rate, moulding, liquor status)

TABLE 2. Facility Characteristics by Study Arm, Kisumu and Meru, Kenya, October 2016 to May 2017

	ePartogram	Paper Partograph	*P* Value
Number of providers, median (IQR)	12 (8)	9.5 (1)	.24
Study-trained providers at each facility, % (IQR)	94.5 (15)	65 (22)	<.001
Volume of deliveries per facility during study period, median (IQR)	101 (84)	76 (224)	.56
Labors recruited into study, % (IQR)	38.5 (57)	43 (46)	.75

Abbreviations: ePartogram, electronic partogram; IQR, interquartile range.

TABLE 3. Client Characteristics by Study Arm, Kisumu and Meru, Kenya, October 2016 to May 2017

	Clients Using ePartogram (n=842)	Clients Using Paper Partograph (n=1,042)	*P* Value
BEmONC facilities, No. (%)	300 (35.6)	317 (30.4)	.02
CEmONC facilities, No. (%)	542 (64.4)	725 (69.6)	.02
Public facilities, No. (%)	842 (100.0)	715 (68.6)	<.001
Faith-based facilities, No. (%)	0 (0.0)	327 (31.4)	<.001
During strike months, No. (%)	439 (53.9)	352 (34.1)	<.001
During non-strike months, No. (%)	376 (46.1)	439 (53.9)	<.001
Parity 0, No. (%)	297 (36.4)	345 (36.6)	.90
Parity 1+, No. (%)	519 (63.6)	642 (63.4)	.90
Admitted during day, No. (%)[a]	496 (58.9)	549 (52.8)	.01
Admitted during night, No. (%)[b]	346 (41.1)	493 (47.3)	.01
Delivered during day, No. (%)[a]	433 (53.1)	542 (52.7)	.80
Delivered during night, No. (%)[b]	382 (46.9)	486 (47.3)	.80
Cervical dilation ≤5 cm at admission, No. (%)	218 (25.9)	371 (36.2)	<.001
Cervical dilation >5 cm at admission, No. (%)	624 (74.1)	642 (63.4)	<.001
Duration of first stage recorded, median (IQR)	120 (180)	555 (360)	<.001

Abbreviations: BEmONC, basic emergency obstetric and newborn care; CEmONC, comprehensive emergency obstetric and newborn care; ePartogram, electronic partogram; IQR, interquartile range.
[a] Day: 6 AM to 6 PM.
[b] Night: 6 PM to 6 AM.

TABLE 4. Clinical Rules Triggered Based on Recorded Parameters Across Type of Partograph Used

	Clients Using ePartogram, No. (%) (n=842)	Clients Using Paper Partograph, No. (%) (n=1,042)	*P* Value
Selected measures of fetal well-being triggered (fetal heart rate, moulding, liquor status)	43 (5.1%)	80 (7.7%)	.29
Selected measures of maternal well-being triggered (pulse, temperature, blood pressure, urine protein)	242 (28.7%)	338 (32.4%)	.49
Selected measures of labor progress triggered (duration and frequency of contractions)	191 (22.7%)	565 (54.2%)	.01

Abbreviation: ePartogram, electronic partogram.

were triggered in 43 (5.1%) of ePartograms and seen on 80 (7.7%) of paper partographs; this difference was not significant. Similarly, selected measures of maternal well-being (maternal pulse, temperature, blood pressure, urine protein) were triggered in 242 (28.7%) of ePartogram users, and noted in 338 (32.4%) of paper partographs; again this difference was not significant. However, clinical rules for abnormal labor progress (duration and frequency of contractions) were triggered in 565 (54.2%) of clients in the paper partograph group compared to 191 (22.7%) of clients in the ePartogram group. In this analysis, we excluded clinical rules related to the alert or the action line as those features of the WHO partograph were eliminated in 2018 WHO guidelines. This difference was statistically significant and has been controlled for in the regression analysis.

Table 5 compares interventions undertaken by providers among clients in the paper partograph and ePartogram groups. Overall, interventions to maintain fetal well-being (position change, oxygen, vacuum extraction, cesarean delivery, or referral) were performed in 116 (14.7%) women in the ePartogram group compared to 54 (5.3%) women in the paper partograph group. This difference was highly significant even when adjusting for differences in the study groups, including effects of the strike period, whether a facility was private, offered CEmONC, the number of providers, parity, and status of cervical dilation at admission (adjusted relative risk [RR]=4.00, 95% CI=1.95–8.19).

Clients in the ePartogram group were more likely than those in the paper partograph group to have been fed during labor (71.3% versus 44.6%, respectively) and given fluids (79% versus 46.8%, respectively), and these differences were significant even when adjusting for the study group differences listed above. In addition, ePartogram clients were encouraged to ambulate and given active management of the third stage of labor more frequently than clients managed with the paper partograph (82% versus 63.3%, respectively) and (92.8% versus 78.7%, respectively), but this difference could not be evaluated for differences in study samples as adjustment models do not converge. Presence of a companion at birth and position of choice, although more practiced in the ePartogram group than the paper partograph group, were not significantly different when adjusting for group differences. Interventions to maintain normal labor (any of ambulation, food, fluids, companion, and position of choice combined) occurred in 778 (86.8%) clients in the ePartogram group compared to 679 (66.9%) clients in the paper partograph group. However, this difference was not statistically significant when adjusting for the factors listed above.

Interventions generally used to address any abnormal progress of labor (augmentation, cesarean delivery) occurred in 31% of the ePartogram group versus 16.7% of the paper partograph group (adjusted RR=1.31, 95% CI=0.67–2.57), but these differences were not significant. The difference remains nonsignificant when we restricted the analysis to only CEmONC sites (two-thirds of study population). The combined rate of labor augmentation and cesarean delivery was 9.4% in the paper partograph users and 10.1% in ePartogram users. When we restricted analysis to CEmONC sites and only those in which any clinical rule was triggered, the combined augmentation/cesarean delivery rates were 44.4% in the paper partograph group and 30.1% in the ePartogram group.

TABLE 5. Actions Undertaken During Labor by Provider Across Type of Partograph Used

Intervention	ePartogram, No. (%[a]) (n=842)	Paper Partograph, No. (%[a]) (n=1,042)	Crude Relative Risk (95% CI)	Crude *P* Value	Adjusted Relative Risk (95% CI)[b]	Adjusted *P* Value
Walk, ambulate	645 (82)	648 (63.3)	1.29 (0.98–1.71)	.07	*Model does not converge*	
Food given	561 (71.3)	456 (44.6)	1.60 (1.14–2.24)	.01	1.73 (1.30–2.30)	<.001
Fluids given	622 (79)	479 (46.8)	1.69 (1.22–2.34)	<.001	1.57 (1.21–2.03)	<.001
Companion present	400 (50.8)	435 (42.5)	1.20 (0.80–1.79)	.39	1.1 (0.70–1.75)	.67
Position of choice	435 (55.3)	255 (24.9)	2.22 (1.05–4.70)	.04	1.49 (0.67–3.30)	.33
Active management of third stage	778 (92.8)	820 (78.7)	1.17 (1.03–1.34)	.02	*Model does not converge*	
Interventions to maintain normal labor	731 (86.8)	697 (66.9)	1.30 (0.96–1.76)	.10	1.09 (0.92–1.29)	.34
Interventions to address fetal well-being	116 (14.7)	54 (5.3)	2.79 (1.03–7.57)	.04	4.00 (1.95–8.19)	<.001
Interventions to address labor abnormality	246 (31)	171 (16.7)	1.86 (0.98–3.55)	.06	1.31 (0.67–2.57)	.42

Abbreviation: CI, confidence interval; ePartogram, electronic partogram.
[a] Percentages exclude missing cases.
[b] Relative risk adjusted for strike period, CEmONC facility, faith-based facility, number of providers, multiparity, and admission after 5 cm dilatation.

In Table 6, we explore the effect of ePartogram use versus paper partograph on the secondary outcome of the study, suboptimal maternal outcomes using a log binomial model.

For suboptimal maternal outcomes, there was no significant difference between the 2 groups after adjusting for facility variables, including disruptions due to the health worker strike, public or private facility, number of providers available at the facility, and client characteristics (parity and cervical dilation on arrival at facility) (adjusted IRR=0.67, 95% CI=0.36–1.26). However, when we restricted analysis only to facilities that offered CEmONC services, we found a significantly lower risk of suboptimal maternal outcomes in the ePartogram group (adjusted IRR=0.15, 95% CI=0.13–0.19).

Table 7 shows the full regression model for the intervention effect of using the ePartogram versus paper partograph for the study's primary outcome of suboptimal fetal outcomes (even when coexisting with other maternal and labor suboptimal outcomes), after adjusting for facility variables, including disruptions due to the health worker strike, whether it was a CEmONC or BEmONC facility, public or private, number of providers available at the facility, and client characteristics (parity and cervical dilation on arrival at facility and selected measures of fetal and maternal well-being and progress of labor). Overall, the ePartogram group had an adjusted IRR of 0.44 (95% CI=0.27–0.73) for adverse fetal outcomes. In other words, ePartogram use was associated with a 56% (95% CI=27%–73%) lower

TABLE 6. Effectiveness of Electronic Partogram on Maternal Outcomes Using Log Binomial Models (n=1,457)

	Crude Relative Risk		Adjusted Relative Risk	
	IRR (95% CI)	*P* Value	Adjusted IRR (95% CI)	*P* Value
ePartogram group	0.81 (0.51–1.28)	.36	0.67 (0.36–1.26)	.21
Strike time	0.99 (0.71–1.39)	.96	1.14 (0.90–1.43)	.27
CEmONC facility	1.13 (0.78–1.63)	.52	0.94 (0.61–1.46)	.80
Faith-based facility	1.39 (1.00–1.94)	.05	1.35 (0.94–1.92)	.10
Number of providers	1.00 (0.96–1.03)	.94	1.03 (0.98–1.08)	.28
Multiparity	0.89 (0.76–1.05)	.17	0.91 (0.78–1.06)	.21
Dilation >5 cm	0.90 (0.80–1.01)	.08	0.95 (0.82–1.09)	.44

Abbreviations: CEmONC, comprehensive emergency obstetric and newborn care; CI, confidence interval; ePartogram, electronic partogram; IRR, incident rate ratio.

TABLE 7. Effectiveness of Electronic Partogram on Fetal Outcomes Using Log Binomial Models (n=1,498)

	Crude Relative Risk		Adjusted Relative Risk	
	IRR (95% CI)	*P* Value	Adjusted IRR (95% CI)	*P* Value
ePartogram group	0.34 (0.22–0.54)	.001	0.44 (0.27–0.73)	.01
Strike time	1.09 (0.77–1.53)	.62	1.33 (0.95–1.88)	.10
CEmONC facility	0.94 (0.45–2.00)	.88	1.13 (0.81–1.57)	.47
Faith-based facility	1.41 (0.83–2.41)	.20	0.80 (0.67–0.95)	.01
Number of providers	0.93 (0.90–0.95)	.001	0.96 (0.93–1.00)	.03
Multiparity	0.67 (0.51–0.88)	.001	0.65 (0.50–0.84)	.01
Dilation >5 cm	0.89 (0.72–1.11)	.31	0.95 (0.82–1.11)	.54

Abbreviations: CEmONC, comprehensive emergency obstetric and newborn care; CI, confidence interval; ePartogram, electronic partogram; IRR, incident rate ratio.

ePartogram use was associated with a 56% lower likelihood of suboptimal fetal outcomes than the paper partograph.

likelihood of suboptimal fetal outcomes than use of the paper partograph, after adjusting for facility and client-level variables.

Because we allocated the 2 largest CEmONC facilities to the ePartogram group and randomly assigned the other 10 facilities to intervention or control, we ran the full regression model restricting only to the 10 randomly assigned sites (4 BEmONC facilities for ePartogram, 4 CEmONC and 2 BEmONC facilities for paper partograph). On this subset, the ePartogram group had an adjusted RR of 0.33 (95% CI=0.22–0.52) for suboptimal fetal outcomes. In other words, ePartogram use in smaller facilities was associated with a 67% (95% CI=48%–78%) lower likelihood of suboptimal fetal outcomes than use of the paper partograph, after adjusting for facility and client-level variables, confirming an even greater effect size than in the total sample.

In Table 8, we present the only before and during study comparison of ePartogram and paper partograph sites for early perinatal mortality estimates based on monthly facility register data during the 6 months before the study compared to facility data during the study period. The outcome of interest (ratio of stillbirths and neonatal deaths in the first 24 hours to total births) is referred to here as early perinatal mortality (EPM). All the EPM recorded in facility registers during the 6 months before the study and during the study regardless of study recruitment status are analyzed here. In ePartogram facilities, the incidence of EPM per 100 births decreased from 2.21 (95% CI=1.77–2.72) in the pre-intervention period to 1.82 (95% CI=1.49–2.20) during the intervention period, giving an adjusted IRR of 0.81 (95% CI=0.66–1.01; *P*=.06). In paper partograph sites, the incidence of EPM per 100 births increased from 2.82 (95% CI=1.99–3.99) before the study to 3.05 (95% CI=2.29–4.05) during the study, giving an adjusted IRR of 1.12 (95% CI=0.89–1.41; *P*=.34).

The ratio of IRRs comparing the change in rates of EPM in ePartogram sites to paper partograph sites before and after implementation, after adjusting for baseline facility characteristics, was 0.73 (95% CI=0.53–1.00; *P*=.049). This suggests that the reduction of EPM in ePartogram sites is trending in the favorable direction compared to paper partograph sites.

Table 9 explores compliance to a set standard for frequency of observations as recorded on the ePartogram or paper partograph. Examination of every 30-minute period of recorded observations showed that the ePartogram recorded all measurements, except for fetal heart rate and contractions, more frequently than the paper partogram. Of the 3 measurements to be recorded every 30 minutes (fetal heart rate, contractions, maternal pulse), maternal pulse was more likely to be recorded on the ePartogram than the paper partograph. All urine measurements (volume, acetone, protein) were recorded more frequently on the ePartogram.

In terms of compliance to measurement frequency norms, only a small fraction of each group was fully compliant for measures to be done every 30 minutes. ePartogram users were more compliant than paper partograph users in measuring pulse, temperature, amniotic fluid status, moulding, blood pressure, and urine; both groups were equally compliant in recording cervical dilation

TABLE 8. Effect of the Electronic Partogram Intervention on Incidence of Early Perinatal Mortality[a] Using Facility Data

	EPM Incidence per 100 Births[b] (95% CI)	EPM Incidence per 100 Births[c] (95% CI)	Adjusted IRR[d] (95% CI)	Difference-in-Difference (95% CI)	*P* Value
ePartogram	2.21 (1.77–2.77)	1.82 (1.49–2.20)	0.81 (0.66–1.01)	0.73 (0.53–1.00)	*P*<.05
Paper partogram	2.82 (1.99–3.99)	3.05 (2.29–4.05)	1.12 (0.89–1.41)		

Abbreviations: CI, confidence interval; ePartogram, electronic partogram; EPM, early perinatal mortality; IRR, incident rate ratio.
[a] Early perinatal mortality comprises neonatal (<24 hours) deaths + fresh stillbirths. Births comprise live births + fresh stillbirths. This analysis included all EPM recorded in the health facility, regardless of whether enrolled in study or not. The model accounts for facility-level clustering and adjusts for facility variables of affiliation (public versus private), capacity to address emergencies (CEmONC versus BEmONC), presence of a medical or clinical officer (yes/no), percentage of providers trained in labor management by study staff (≥75% or <75%), median years of provider experience, and whether a health worker strike was occurring in the month in public facilities (yes/no).
[b] Six-month period before study.
[c] During study implementation.
[d] Comparing before and during study rates in each group.

TABLE 9. Compliance With Recommended Frequency of Recording of Observations During Labor

Measure	Instances When Observation Could Be Recorded, %		Instances When Observation Should Have Been Recorded Per Standard, %	
	ePartogram	Paper Partograph	ePartogram	Paper Partograph
Recommended timing:	**Every 30 minutes**		**At least every 1 hour**	
Fetal heart rate	76.0	84.0	80.2	85.9
Contractions	76.0	82.0	79.5	90.7
Maternal pulse	72.0	42.0	73.9	33.8
Recommended timing:	**Every 2 hours**		**At least every 2 hours**	
Temperature	43.0	18.0	99.3	72.6
Recommended timing:	**Every 4 hours**		**At least every 4 hours**	
Color of amniotic fluid	43.0	23.0	100.0	65.8
Moulding	42.0	18.0	100.0	60.1
Cervical dilatation	51.0	35.0	100.0	99.8
Descent	50.0	34.0	99.7	98.0
Blood pressure	42.0	25.0	99.7	92.2
Urine protein	32.0	5.0	95.8	29.0
Urine acetone	33.0	5.0	96.1	28.5
Urine volume	33.0	5.0	96.4	31.6

Abbreviation: ePartogram, electronic partogram.

and descent of the presenting part; and paper partograph users were more compliant with recording fetal heart rate and contractions.

Qualitative Interviews

Feasibility and acceptability of the ePartogram technology were examined during in-depth interviews. SBAs noted minimal issues with care and maintenance of the tablet and were able to keep the devices clean. Some electricity outages were noted, but SBAs reported no issues with charging tablet devices.

SBAs unanimously agreed that once they were trained with the ePartogram, data entry was simple and user-friendly. Several SBAs cited issues with the inability to correct ePartogram entry errors after 15 minutes had elapsed. If a parameter was entered incorrectly and not caught in time, the system would continually alert the SBA to a nonexistent concern. The SBAs found this distracting.

SBAs also noted challenges with ePartogram data entry when caring for multiple clients.

> *Like maybe when you are entering the data for a mother who has just come, in regard to the time you are entering the fetal heart, the contractions, the dilations so after you are through with the mother and now you want to enter [data], you find there have a difference of 1 minute, 1 minute, 1 minute so assuming you have 3 or 4 mothers then it will reach the due time for entering the next data; that difference for 1 minute for 3 or 4 mothers, it is hectic.* —SBA at a health center

The ePartogram emits audible reminders when it is time for an SBA to take a clinical measurement. Inputting an abnormal clinical value triggers an alarm. SBAs noted that the reminders were clear and helpful, and they especially appreciated that the ePartogram dashboard within the application highlighted when to take measurements and which measurements were overdue. In-charges also noted that the ePartogram improved real-time measurement recording and reduced retrospective data entry.[17] SBAs reported that having reminders increased the frequency with which they checked clients and prompted them to take clinical action when needed compared to use of the paper partograph.

> *For me, it is just the quality of work, because for the paper partograph, you just leave it . . . until when after the*

mother delivers you fill in what you had not filled in, but for (the ePartogram) you have to be on toes so that if there is a problem or some necessary intervention, it just shows you immediately instead of waiting for a delay that maybe will result to other unwanted things like fetal distress, maternal distress. —SBA at a health center

SBAs made clear that the ePartogram was not a panacea for all clinical decision making challenges and that attitudes and behavior change must also be addressed.

[T]he device will only give you an alert and if you are asleep or you are not acting on what next needs to be done, the alerts will continue and the outcome will be negative, so if we also need behavior change, that would be good. —Hospital in-charge

Alerts also prompted action for urinalysis for protein.

There are some of the things that have not been done routinely previously. But currently, the device is helping us to do, like, the urine test. . . . [N]owadays every client has to be tested. Before we would do the urine test only when we realize the blood pressure is elevated, but today the device will not allow you to continue. It will keep on ringing [an] alarm, so you have no option but to do it. —Hospital in-charge

SBAs were pleased with the ePartogram but raised concerns about using it at high-volume sites because timely data entry with more clients might be difficult.

SBAs were overwhelmingly pleased with the technology, but raised concerns about introducing the ePartogram at high-volume sites. SBAs often noted that once the facility reached around 4 clients per SBA, timely data entry was difficult. Additionally, they suggested that other team members be trained to use the ePartogram (e.g., surgical staff who take over care once a client gets to the operating room, as well as students since they often monitor clients). SBAs were concerned that failure to act or respond to an alert due to staff shortage or more pressing complications with other clients may be recorded in the ePartogram and viewed as neglect. However, despite the challenges, SBAs agreed that the ePartogram was a great improvement over the paper partograph, even at high-volume sites.

The ePartogram's use was associated with improved fetal outcomes compared to paper partographs, and the intervention was well accepted by care providers and their supervisors.

Because my experience is that ePartogram helps us to manage labor better. In busy facilities, it becomes difficult to enter all that information on the paper partograph because paper takes a lot of time before you fill, but it is easier with ePartogram. —SBA at a health center

Supervisors noted that ePartogram improved their ability to oversee both their subordinates and patients at both a clinical and management level.

The supervisor application is good in such a way that I have it all the time, I can monitor what work is going on in the labor ward if there is something that needs my attention. I could call the nurse on duty and she would check and see so the supervisor's application is there to support the nurse on the ground. —Physician supervisor at a large hospital

The ePartogram provides real-time sharing of labor observations, enabling on-site and off-site supervisors, as well as SBAs at referral facilities, to view client information and provide consultation as needed. Supervisors noted that this feature improved their ability to oversee their subordinates and clients on both a clinical and managerial level. They also noted that the ePartogram improved communication in both directions and that SBAs used the supervisory communication system to voice concerns about clients and ask questions.

[W]henever there is a need I can always come very fast from wherever I am. I am aware what is happening in [the] labor ward. If I am outside the hospital, I am aware [of] what is happening in [the] labor ward and we have kept that continuous conversion. —Hospital in-charge

For SBAs in referral facilities, the ePartogram allowed surgical teams to better prepare for incoming referral clients and to communicate in a network. Although most facilities communicated by phone during referrals when using the paper partograph, the ePartogram provided a more detailed account of the client's labor and complications. One SBA working in a large facility hypothesized that a shift in the number of cesarean deliveries in her facility was due to early identification and management of complications.

Yes, when we are using ePartogram in our labor ward, the C-section rate goes down, the number of complications goes down because they are doing early interventions and you are able to bring everybody on board so that now, it is a team work approach, the picture is better. I am now imagining [if] other facilities are using it, we will actually cause a paradigm shift about maternal health care in the system. —Hospital in-charge

DISCUSSION

Our study is one of the first to test the feasibility and effectiveness of using an electronic decision-support tool for intrapartum care in limited-resource settings. The results of this quasi-experimental study demonstrated the effectiveness of SBA use of an electronic partograph on an Android tablet. The ePartogram's use was associated with

improved fetal outcomes compared to use of paper partographs, and the intervention was well accepted by care providers and their supervisors. Furthermore, use of the ePartogram was associated with a trend in a decline in EPM after the introduction of ePartogram compared to the preceding 6 months when the facilities used the paper partograph. To our knowledge, this is the first report on the effectiveness of a digital labor decision-support system in limited-resource settings.

In Kenya, as in most limited-resource countries, SBAs monitor labor using the paper partograph while providing care to clients to maintain normal labor, manage complications, and assist delivery. Often, the task of monitoring labor is delegated to junior midwives or students, with senior SBAs checking in periodically. Ideally, monitoring fetal and maternal well-being and labor progress will enable SBAs to identify or predict complications immediately and take appropriate actions, resulting in improved maternal and fetal outcomes. In our study, we reinforced labor monitoring and good routine care norms and protocols in both groups during a 2-day refresher training but did not attempt to judge whether the actions taken during the study were appropriate to the problem identified by the monitoring process. For those using the ePartogram, the additional input was entering observations in the tablet device that gave audible alerts if a measurement was not taken on time and visual alerts when a single or a combination of more observations were determined to be abnormal or predicted a potential complication. SBAs valued the ePartogram's audible and visual alerts, which allowed them to spend more time determining what might be wrong and why, rather than if something was wrong.

Our study found that the ePartogram group had lower rates of clinical rules triggered, significantly so for duration and frequency of contractions. One explanation is that with an ePartogram, an audible sound alerts the provider and can also be seen by the supervisor, so actions are more likely to be promptly taken to correct the situation. With the paper partograph, a provider may fail to recognize that a clinical rule has been triggered and actions may be delayed, potentially setting off another clinical trigger. In-depth interviews with supervisors indicated that they were proactive when they are also being alerted in real time by the app rather than awaiting to be called, as with the paper partograph. This would be an important area to research further. One concern is that providers may get alarm fatigue and choose to ignore auditory alerts.

As technology use in labor rooms increases, there is some concern that personal attention and care will suffer.[29] Our study showed that clients in the ePartogram group were more likely to receive care to maintain normal labor, including encouragement to ambulate, feed, drink fluids, and receive active management of the third stage of labor.

One challenge with the paper partograph is that it is sometimes used solely for recordkeeping rather than decision making, and labor measurements are often entered retrospectively, which the ePartogram does not allow.[11–14,17] Although our study showed that fetal heart rate and contraction measurements were recorded more frequently in the paper partograph group, we did not evaluate whether or not monitoring frequency was exaggerated in some cases or if some of these were entered retrospectively.

It was noteworthy that users of the ePartogram found that taking care of 4 or more patients limits the feasibility of use or usefulness of the ePartogram. We do not believe that the ePartogram should be used to take care of more patients than used for paper-based labor charts. However, time-saving features, such a drop-down menus and auto plotting, make it easier to do the right thing and providers have more time to provide respectful care, perform interventions, and perform observations on time, all of which would improve outcomes.

Interviews confirmed the ease of entry with the ePartogram, and ePartogram users appreciated the drop-down menus, as also noted by Litwin et al.[22] Further, interviewed in-charges and SBAs stressed that the ePartogram prompted action, and they discussed specific examples of the ease of use during labor management and decisions to refer the laboring client or operate.

The observation by a hospital supervisor that he had observed lower cesarean delivery rates warrants further analysis, but our data indicate that at CEmONC sites and among patients who had any clinical rule triggered, the combined labor augmentation/cesarean deliveries are lower in ePartogram-managed labors, but our study was not powered sufficiently to attribute this to ePartogram use. However, we found that suboptimal maternal outcomes were significantly lower when the ePartogram was used in CEmONC facilities.

Unfortunately, the rate of cervical dilation has dominated decision making instead of using the combination of measures of fetal and maternal well-being and progress of labor, all of which are visually represented in the paper partograph. The ePartogram depends on clinical rules that place

emphasis on all these measurements, not just alert and action lines. The latest WHO guidelines and many clinicians have questioned Friedman's labor curve. Of the 77 clinical rules that form the decision-support part of the ePartogram, only 3 are related to the rate of progression of cervical dilation and the alert/action line. Another 4 are related to the action line but in relation to number and duration of contractions and descent of presenting part.

The ePartogram is based on clinical rules that were developed and validated between 2013 and 2016 using WHO's guidance[30] at that time and clinical consultation. Since then, WHO completed a large prospective cohort study[31] and has issued new recommendations for intrapartum care for a positive childbirth experience.[32] By far the most consequential new recommendations relate to definitions of onset of the first stage of labor. First stage is now defined as regular painful contractions, substantial effacement, and more rapid dilation from 5 cm. The active phase usually does not exceed 12 hours in first labors and 10 hours in subsequent labors. In spontaneous labor, cervical dilatation of 1 cm per hour during active phase, as depicted by the alert line, does not accurately identify women at risk of adverse birth outcomes, and a minimum cervical dilation of 1 cm per hour is unrealistic for some women. WHO is developing a new partogram; however, its adoption will likely require a massive effort, including efforts to replace the existing paper partograph. The ePartogram's clinical decision-support system consists of 77 clinical rules, 8 of which will be changed as a result of the new guidelines. We believe that a digitally connected app like the ePartogram can accelerate adoption of the new guidelines at scale.

Challenges of introducing a digital labor device include cost, cybersecurity, and equipment malfunction.

In 1955, Friedman published the labor progression curve that drove obstetric practice for more than 40 years.[33] Perhaps less known is that his conclusions were based on a study involving only 500 women with an average age of 20 years. Fifty-five percent had forceps delivery; 9% underwent cesarean delivery; and 22% received "twilight sleep," 42% with moderate sedation and 31% with heavy sedation.[33] By no stretch of the imagination can we call that a "normal" population to define normal labor or create labor cervicographs. Yet we continued to use Friedman's curve well into this century. In 2002, Zhang reassessed the labor curve in nulliparous women,[34] and in 2013, Boyle et al. examined 38,484 first-time cesarean deliveries, of which 30.8% were done for primigravidas, 31% were due to "failure to progress," and 40% were for women less than 5 cm dilated.[35] They concluded that more than 10% of primigravidas have unplanned and unnecessary cesarean deliveries due to "failed progress" in early labor. These studies have been confirmed in Asia, Latin America, and most recently in Africa.[31] The ePartogram digitizes labor data and holds it in a central server. Digitizing labor measurements and including prenatal, outcome, and postnatal data offers the exciting possibility of harnessing artificial intelligence and machine learning principles to further refine clinical rules and enhance clinical decision-support in a manner that has never been possible before without relying on sample studies.[36,37]

In our analysis of measures of labor progress, we excluded cervical dilation and only included number and duration of contractions and descent of presenting art, in recognition of new WHO guidelines and the many studies that confirm that Friedman's original assertion of 1 cm per hour is flawed. One additional value of a digital labor chart is in our ability to add or eliminate rules as they are formulated or as new evidence dictates.

The challenges of introducing a digital labor device include cost, cybersecurity, and equipment malfunction. Although the unit cost of tablets has declined, other cost implications exist. Smartphones are ubiquitous and less expensive. Furthermore, clinical workflows may need to be adapted to ensure that the ePartogram is implemented efficiently. For example, in the larger facilities, SBAs teach students how to use the paper partograph, so it was challenging for them to teach the paper partograph while using the ePartogram on their clients. Against these challenges were the ability of supervisors to access labor data in real time and the existence of complete digital records. Therefore, a cost-benefit analysis is needed. A major challenge to overcome is building a sufficiently robust central server to host the vast amount of labor data being generated in real time.

An adapted version of the ePartogram is being tested in India on a larger scale. It uses prevailing labor charts in India and also digitizes other pregnancy data and outcome data. We hope to align this version with new WHO guidance by altering 8 of the 77 clinical rules and using the newly developed WHO labor charts that are undergoing testing right now. In addition, secondary analysis of data will reveal what proportion of interventions are warranted and not warranted by adjusting these clinical rules. We believe future research should focus on rationalizing the clinical rules so as to prevent alert fatigue and evaluate with more

robust design the supervisor function. An important area to explore further is compliance to set standards of observations especially when interventions, such as augmentation, are performed. Finally, there is a potential to harness the growing volume of digitized labor, intervention, and outcome data for machine learning and artificial intelligence.

Study Limitations

An important limitation to our study is that the ePartogram group and paper partograph group were not similar in a number of key characteristics, including the fact that the 2 largest public facilities were in the ePartogram group and the 2 faith-based hospitals were in the paper partograph group. The study was disrupted by a prolonged health worker strike, the effect of which may have been unevenly distributed among sites and probably led to a greater proportion of patients arriving much later in labor than we had expected. We have addressed these differences by adjusting for them in all outcome analyses.

One design limitation was that the 2 largest facilities were allocated to the ePartogram arm and the other 10 smaller facilities were randomly distributed. Our analysis of only these 10 randomly assigned sites showed an even bigger effect size for the primary outcome of suboptimal fetal outcome than when we included all 12 facilities in our analysis.

One limitation was that a greater proportion of patients in the ePartogram group came after 5 cm dilation than in the paper partograph group and that median duration of labor recorded on paper partographs was much longer than on the ePartogram. One possible explanation is that the ePartogram did not allow for retrospective entry of data after a certain period of time (30 minutes after the parameter recording was due) but such entries were possible on the paper version. Another possible explanation is that providers may be waiting longer to start the ePartogram than with paper partographs, but this was not raised during our in-depth interviews. We do, however, show that while there was a longer period of labor observed in the paper partograph, there was much poorer compliance to labor observation norms.

The study sample facilities are heterogeneous, with different levels of care and management, and the distribution across the ePartogram and paper partograph groups differed.

The information recorded was not always complete and differed by ePartogram and paper partograph groups, possibly resulting in a differential bias across the groups. In Table 8, we show that except for 2 parameters (fetal heart rate and contractions,), every other observation is better recorded in the ePartogram group, and there is, in fact, less missing recording of observations in the ePartogram, perhaps as a result of the "nag" feature that alerts a provider and their supervisors that observation are overdue. The data collected as part of the study are entirely representative of the routine status of patient records. Methods to assess missingness assume that the missing data are random or one is able to understand and explain the pattern. Hence, we chose not to use any of these methods, as they would obfuscate the data.

SBAs in the study may have based their decision to start a client on the ePartogram, paper partograph, or not start a partogram at all on criteria such as staffing and number of clients under observation for labor. Many clients probably arrived in very late stages of labor so there was not time to record more than 1 set of parameters (at admission). We are unable to completely explain the processes that went into which woman was put on the partographs and what data were entered. These conditions, which differed by facility, may have resulted in a selection bias in the sample of women, and this bias may be difficult to quantify in direction or magnitude. Thus, the ePartogram was implemented within a clinical setting where both the ePartogram and paper partographs were used, and during this introductory phase, not all clinical workflows were updated to accommodate ePartogram use.

An additional limitation is interpreting the reduction in early perinatal mortality in ePartogram sites because we provided refresher training to a greater number of SBAs in those sites than in paper partograph sites even though we have adjusted for that difference. Also, our training of SBAs that included reinforcing recordkeeping occurred during the 6 months before the intervention, and it is possible that records were better in the period after training than in the pre-intervention period. We do not think this limits conclusions from our other analysis as we only included labor charts from study-trained SBAs for all outcomes.

Data collection was delayed (particularly in public-sector facilities) due to a health worker strike in both Meru and Kisumu counties. The ePartogram application performance and speed

slowed after 1 to 2 months of clinical use, and updated software was deployed to address these challenges. The study's main conclusion of improvement in fetal outcomes among ePartogram users remains significant even after adjusting for facility characteristics.

The differences between the groups may (and would) have resulted in the observation of differences in the outcome. However, not all the effect size can be explained by the baseline differences alone. The intervention group may have been subjected to greater scrutiny and more frequent supervision due to the use of a novel technology. The novelty of tablets alone is likely to improve the attention of providers and contribute to better outcomes. Given the drawbacks in the design of the study, we feel that the effect size should not be the only evidence of impact but placed in the context of the rest of the data.

CONCLUSION

Use of the ePartogram resulted in significant improvements in fetal outcomes and use of interventions to maintain normal labor compared to the paper partograph. SBAs using ePartograms were also at least as or more compliant in recording measurements during labor compared to SBAs using paper partographs—they noted that the alerts prompted them to take clinical action. As WHO develops a revised partograph that reflects its 2018 intrapartum care recommendations, the ePartogram has great potential to improve quality of care and outcomes through consistent and meaningful labor monitoring and clinical care.

Acknowledgments: The authors acknowledge Timothy Muhavi for assistance with study implementation and data management, India Richter for data quality assurance, Blair Olivia Berger for assistance with data analysis, D-tree International for app development support, and Jean Sack for support on literature reviews. During the study period, Dr. Dickens Onyongo, formerly the Director of Health, Kisumu County, and Dr. James Gitonga, formerly the Chief Officer for Health, Meru County, approved the study and provided strategic oversight of the study design and implementation in Kisumu and Meru, respectively. Instrumental to this study were the master trainers and Ministry of Health representatives who trained study participants in the intervention and control groups and contributed to study implementation: Jane Awuor, Pamela Jacinda, Simon Munyoki, and Dennis Kiambi, Meru; Rhoda Njeri, and Zipporah Murethi (Jhpiego). The health resource information officers at each facility provided support in de-identifying data records and scanning them into the study database. We thank the following individuals for support with data extraction: University of Nairobi (UON) midwives Raheli Misiko Mukhwana, Anne Khaoya Nendela, and Edith Waithira Gicheha; UON obstetrician-gynecologists Alfred Osoti, George Gwako, and Diana Ondieki; Johns Hopkins Bloomberg School of Public Health students Ameneh Amini, Mindi DePaola, Priyam Doshi, Elizabeth Kang, Victoria Laney, Kelly Li, Nishit Patel, Rahul Rajkumar, and Rebecca White; and interns in Kenya, Mercy Cherotich and Rahab Ngata. We thank Joan Taylor for copyediting a previous draft of this article.

Funding: The study was funded by Jhpiego, an affiliate of Johns Hopkins University, and GE Foundation. Ongoing ePartogram app adaptation and maintenance was funded by GE Foundation and Laerdal Foundation. The GE Foundation and Laerdal Foundation had no role in study design, data collection, data analysis, data interpretation, or manuscript development.

Competing interests: None declared.

REFERENCES

1. World Health Organization (WHO), UNICEF, United Nation Population Fund, World Bank Group, United Nations Population Division. *Trends in Maternal Mortality: 1990 to 2015. Estimates by WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division*. https://www.who.int/reproductivehealth/publications/monitoring/maternal-mortality-2015/en/. Geneva: WHO; 2015. Accessed March 28, 2019.
2. Kassebaum NJ, Bertozzi-Villa A, Coggeshall MS, et al. Global, regional, and national levels and causes of maternal mortality during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet*. 2014;384(9947):980–1004. CrossRef. Medline
3. Lawn JE, Blencowe H, Waiswa P, et al; Lancet Ending Preventable Stillbirths Series study group; Lancet Stillbirth Epidemiology investigator group. Stillbirths: rates, risk factors, and acceleration towards 2030. *Lancet*. 2016;387(10018):587–603. CrossRef. Medline
4. Lawn JE, Lee AC, Kinney M, et al. Two million intrapartum-related stillbirths and neonatal deaths: Where, why, and what can be done? *Int J Gynaecol Obstet*. 2009;107(Supplement):S5–S18, S19. CrossRef. Medline
5. Chou D, Daelmans B, Jolivet RR, et al. Ending preventable maternal and newborn mortality and stillbirths. *BMJ*. 2015;351:h4255. CrossRef. Medline
6. Kieny M-P, Evans TG, Scarpetta S, et al. *Delivering Quality Health Services: A Global Imperative for Universal Health Coverage*. Washington, DC: World Bank Group; 2018. http://documents.worldbank.org/curated/en/482771530290792652/Delivering-quality-health-services-a-global-imperative-for-universal-health-coverage. Accessed March 26, 2019.
7. Tunçalp Ö, Were WM, MacLennan C, et al. Quality of care for pregnant women and newborns-the WHO vision. *BJOG*. 2015;122(8):1045–1049. CrossRef. Medline
8. World Health Organization (WHO). *WHO Recommendations for Augmentation of Labor*. https://www.who.int/reproductivehealth/publications/maternal_perinatal_health/augmentation-labour/en/. Geneva: WHO; 2014. Accessed March 26, 2019.
9. Windrim R, Seaward PG, Hodnett E, et al. A randomized controlled trial of a bedside partogram in the active management of primiparous labour. *J Obstet Gynaecol Can*. 2007;29(1):27–34. CrossRef. Medline
10. Maternal and Child Health Integrated Program. Quality of Antenatal and Delivery Care Services in Six Countries in Sub-Saharan Africa. https://www.mchip.net/sites/default/files/Quality%20of%20ANC%20in%206%20African%20Countries.pdf. Published 2016. Accessed March 28, 2019.
11. Ollerhead E, Osrin D. Barriers to and incentives for achieving partograph use in obstetric practice in low- and middle-income countries: a systematic review. *BMC Pregnancy Childbirth*. 2014;14(1):281. CrossRef. Medline
12. Asp G, Sandberg J, Ezechi O, Pettersson KO. Challenges of immediate newborn care in maternity units in Lagos, Nigeria: An observational study. *J Obstet Gynaecol*. 2011;31(7):612–616. CrossRef. Medline
13. Mathibe-Neke JM. Facilitation of midwifery students regarding utilisation of a partograph. *Afr J Nurs Midwifery*. 2009;11(1):34–47.

https://pdfs.semanticscholar.org/23da/dd2eb33f00a85d65103ed66486c7fde431ee.pdf.

14. Delvaux T, Aké-Tano O, Gohou-Kouassi V, Bosso P, Colllin S, Ronsmans C. Quality of normal delivery care in Côte d'Ivoire. *Afr J Reprod Health.* 2007;11(1):22–32. CrossRef. Medline
15. Kagema F, Ricca J, Rawlins B, et al. *Quality of Care for Prevention and Management of Common Maternal and Newborn Complications: Findings from a National Health Facility Survey in Kenya.* http://www.mchip.net/sites/default/files/KenyaQoCreportfinal.pdf. Baltimore, MD: MCHIP; 2011. Accessed March 28, 2019.
16. Lavender T, Cuthbert A, Smyth RMD. Effect of partograph use on outcomes for women in spontaneous labour at term and their babies. *Cochrane Database Syst Rev.* 2018;8(8):CD005461. CrossRef. Medline
17. Bedwell C, Levin K, Pett C, Lavender DT. A realist review of the partograph: when and how does it work for labour monitoring? *BMC Pregnancy Childbirth.* 2017;17(1):31. CrossRef. Medline
18. Equipping providers with digital tools to improve maternity care: Project iDeliver. Merck for Mothers website. http://www.merckformothers.com/docs/MERKforMothers-iDeliver-Apr13-R3.pdf. Published April 2018. Accessed March 28, 2019.
19. Dimagi. mLabor (Intrapartum Care): CommCare for Intrapartum Care. https://www.dimagi.com/sectors/intrapartum-care/. Published 2018. Accessed October 10, 2018.
20. Underwood H, Ong'ech J, Omoni G, Wakasiaka S, Sterling SR, Bennett J. Improving partograph training and use in Kenya using the Partopen digital pen system. In: Fernández-Chimeno M, Fernandes PL, Alvarez S, Stacey D, Solé-Casals J, Fred A, Gamboa H, eds. *Biomedical Engineering Systems and Technologies.* 6th International Joint Conference, BIOSTEC 2013. Berlin, Germany: Springer; 2014:407–422.
21. Stuchlik P. *Record, Representation and Use of Electronic Partogram and Cardiotocogram.* [master's thesis]. Prague: Czech Technical University in Prague; 2015.
22. Litwin LE, Maly C, Khamis AR, et al. Use of an electronic partograph: feasibility and acceptability study in Zanzibar, Tanzania. *BMC Pregnancy Childbirth.* 2018;18(1):147. CrossRef. Medline
23. World Health Organization (WHO). Making pregnancy safe: the critical role of the skilled birth attendant. A joint statement by WHO, ICM and FIGO. http://apps.who.int/iris/bitstream/10665/42955/1/9241591692.pdf. Published 2004. Accessed March 26, 2019.
24. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;42(2):377–381. CrossRef. Medline
25. *Stata* [computer program]. Version 14.2. College Station, TX: StataCorp; 2015.
26. Mack N, Woodsong C, MacQueen KM, Guest G, Namey E. *Qualitative Research Methods: A Data Collector's Field Guide.* Research Triangle Park, NC: FHI 360; 2005.
27. *ATLAS.ti* [computer program]. Version 7. Berlin, Germany: GmbH D; 2015.
28. Ritchie J, Lewis J, Nicholls CM, et al. *Qualitative Research Practice: A Guide for Social Science Students and Researchers.* 2nd ed. London, United Kingdom: Sage Publications, Ltd.; 2013.
29. Bailey JE. Does health information technology dehumanize health care? *Virtual Mentor.* 2011;13(3):181–185. CrossRef. Medline
30. World Health Organization (WHO), UNICEF, United Nations Population Fund. *Managing Complications in Pregnancy and Childbirth: A Guide for Midwives and Doctors.* Geneva, Switzerland: WHO; 2003.
31. Oladapo OT, Souza JP, Fawole B, et al. Progression of the first stage of spontaneous labour: a prospective cohort study in two sub-Saharan African countries. *PLoS Med.* 2018;15(1):e1002492. CrossRef. Medline
32. World Health Organization (WHO). *WHO Recommendations: Intrapartum Care for a Positive Childbirth Experience.* Geneva, Switzerland: WHO; 2018. http://www.who.int/reproductivehealth/publications/intrapartum-care-guidelines/en/. Accessed October 4, 2019.
33. Friedman EA. Primigravid labor: a graphicostatistical analysis. *Obstet Gynecol.* 1955;6(6):567–589. CrossRef. Medline
34. Zhang J, Troendle JF, Yancey MK. Reassessing the labor curve in nulliparous women. *Am J Obstet Gynecol.* 2002;187(4):824–828. CrossRef. Medline
35. Boyle A, Reddy UM, Landy HJ, Huang CC, Driggers RW, Laughon SK. Primary cesarean delivery in the United States. *Obstet Gynecol.* 2013;122(1):33–40. CrossRef. Medline
36. Radin JM, Topol EJ, Steinhubl SR. A digital toolkit for improved maternal health. *Lancet.* 2018;392(10153):P1107. CrossRef
37. Desai GS. Artificial intelligence: the future of obstetrics and gynecology. *J Obstet Gynaecol India.* 2018;68(4):326–327. CrossRef. Medline

Peer Reviewed

Received: May 26, 2019; **Accepted:** October 15, 2019

Cite this article as: Sanghvi H, Mohan D, Litwin L, et al. Effectiveness of an electronic partogram: a mixed-method, quasi-experimental study among skilled birth attendants in Kenya. *Glob Health Sci Pract.* 2019;7(4):521-539. https://doi.org/10.9745/GHSP-D-19-00195

GLOBAL HEALTH: SCIENCE AND PRACTICE
Dedicated to what works in global health programs
OPEN ACCESS

ORIGINAL ARTICLE

Provider-Initiated Family Planning Within HIV Services in Malawi: Did Policy Make It Into Practice?

Erin K. McGinn,[a] Laili Irani[b]

Four years after Malawi embraced a policy of provider-initiated family planning (PIFP) within its HIV Clinical Guidelines, this policy remained largely unimplemented at the health facility level. Strengthening PIFP in Malawi's public and private health facilities will require targeted and comprehensive systems changes.

ABSTRACT

Malawi's guidelines for *Clinical Management of HIV in Children and Adults* promote the practice of provider-initiated family planning (PIFP) for all clients over the age of 15. The guidelines recommend that providers should offer all clients condoms, offer injectables to female clients, and refer clients to another provider or site if clients prefer another family planning method. This study assessed to what extent family planning services had been integrated into HIV services among 41 facilities in Malawi (hospitals, health posts, health centers; public and nonprofit private) and how the reproductive rights of people living with HIV were being addressed. Data were collected through facility audits (N=41), provider interviews (N=122), client exit interviews (N=425), and mystery client visits (N=58). This study found that contrary to clinical protocols, only 14% of clients at the antiretroviral therapy (ART) clinic had reported being asked about their family planning/fertility intentions during the visit that day. Only 24% of providers at the facility had received training on family planning-HIV integration, and 21% had no family planning training at all. Overwhelmingly, ART clients relied on condoms to meet their family planning needs. Only 24% of ART clinics had injectables available, and only 15% of ART clinics had a full range of family planning methods (short- and long-acting, hormonal and nonhormonal) available to clients. These findings suggest that Malawi's strong national policies on family planning-HIV integration, and specifically PIFP, are not being implemented in practice and thus not adequately addressing the family planning needs of clients with HIV. To improve PIFP, Malawi requires targeted systems changes. Facilities need to broaden their family planning method mix offerings. Furthermore, providers need more training on family planning and the importance of HIV clients having access to family planning services, and referral services need to be strengthened so providers can ensure clients have access to their method of choice in a timely manner.

INTRODUCTION

What Is Provider-Initiated Family Planning?

Provider-initiated family planning (PIFP) is an approach that encourages health care providers to routinely and proactively ask about a client's reproductive intentions during the client-provider interaction, even if the client has come for other health services. The provider asks the client whether he/she intends to have another child (ever or in the near future) and what they might be using to space or avoid an unintended pregnancy. The provider then counsels the client on family planning and offers and/or refers the client for a range of contraceptive methods (or may provide information or referral for the client's partner/spouse as needed).[1] PIFP is rights-based, and the client can decline counseling and/or contraceptive methods as desired. The purpose of PIFP is to ensure there is "no missed opportunity to offer family planning."[2] Furthermore, a comprehensive review of over 2,500 articles shows that promoting voluntary family planning as part of routine HIV services is the number one evidence-based practice on how to meet the sexual and reproductive health needs (and rights) of women living with HIV.[3]

PIFP mimics the successful strategy of provider-initiated counseling and testing (PICT). In 2007, the World Health Organization established global guidelines on PICT for HIV, which encourages providers to routinely offer HIV testing as part of standard medical care for all patients in the context of a generalized HIV epidemic.[4] PICT takes an "opt-out" approach, in that clients must specifically decline the test. Studies have shown a positive correlation between PICT and an uptake in HIV testing and condom use; therefore, as a programmatic intervention, PICT has been instrumental in increasing the number of people being tested for HIV and using condoms for HIV prevention.[5]

Numerous studies have shown that women and couples with HIV have a high unmet need for either limiting

[a] The Palladium Group, Washington, DC, USA.
[b] Population Council, New York, NY, USA.
Correspondence to Erin McGinn (erin.mcginn@thepalladiumgroup.com).

BOX. Malawi National Clinical Guidelines Instructions for Providers on Provider-Initiated Family Planning

- Assume all patients aged 15 years and older are sexually active.
- Offer condoms to all men and condoms and Depo-Provera [injectables] to all women.
- Give patients the opportunity to refuse either method.
- Refer clients to family planning clinics for further counseling or for other family planning methods.

Source: Section 6.5.1 of Malawi's 2014 *Clinical Management of HIV in Children and Adults*.

childbearing or delaying conception until their health or other personal circumstances improve.[6–8] A recent 2016 study of pregnancy intentions of 220 women with HIV on antiretroviral therapy (ART) in a district hospital in Lilongwe, Malawi, found that 75% of them reported their current pregnancy was unintended (16% mistimed and 59% unwanted).[9] HIV clients, like everyone else, deserve comprehensive family planning counseling and services and access to a full range of contraceptive methods to meet their reproductive intentions and life context, which likely change over time.

Policies and Guidelines in Malawi

Malawi has a high prevalence of HIV. The 2015 Malawi Demographic and Health Survey (DHS) estimated an overall HIV prevalence of 8.8%, with a 10.8% prevalence among women and a 6.4% prevalence among men. Supplement 1 contains details on relevant indicators pulled from the last 2 rounds of the Malawi DHS and the findings from this study. Given this high prevalence and a growing number of patients seeking ART services, providing family planning services within ART clinics can be an effective way to address unmet need for contraception.

In 2011, the Government of Malawi issued its integrated guidelines for providing HIV care in its primary health care services, called *Clinical Management of HIV in Children and Adults*. These guidelines recommended PIFP be practiced at every scheduled visit at pre-ART follow-up and in ART clinics.[10] Providers are instructed to assume all patients 15 years and older are sexually active and to proactively offer (30 or more) condoms to all men and women and injectables to women 15 years and older. The guidelines do take a rights-based approach in acknowledging that all patients with HIV have a right to a full reproductive life and should have the information on all family planning methods and be free to choose any family planning method of their choice. The guidelines further emphasize the need for providers to give clients an opportunity to refuse either method and to refer clients for counseling or other family planning methods (Box). Malawi updated these guidelines in 2014, with no substantial changes to the PIFP section.

The findings reported in this article are part of a larger multipronged assessment on the integration of family planning and HIV services in Malawi, undertaken in 2014 and 2015 by the United States Agency for International Development (USAID)-funded Health Policy Project.[11] This study asked several implementation-related questions to ascertain whether Malawi's clinical guidelines on PIFP were functioning at the facility level, including:

- What was the availability of family planning information and commodities within ART services?
- How are services organized to promote family planning-HIV integration?
  - Are the health workers providing HIV services trained in family planning?
  - For clients who need referrals, do providers know about referral points, i.e., other times or locations to which clients can be referred to obtain additional (family planning) services?
- Do clients on ART need and want family planning services?
- Do providers offer family planning to or initiate family planning with ART clients?

The objective of this article is to highlight the extent to which PIFP was being implemented per government guidelines, describe some system barriers to PIFP implementation, and provide some key recommendations to improve PIFP.

METHODS

Study Setting

Quantitative and qualitative data were collected between April and May 2015 from 41 facilities across 9 districts in Malawi—3 districts from each of the Northern, Central, and Southern Regions (Table 1). The districts were stratified by region and then randomly selected. A purposive sample of public and private facilities was chosen to represent a range of facility types. These included public

TABLE 1. Types of Facilities From Which Data Were Collected Across 9 Districts in Malawi, by Facility Type, April–May 2015

Type of Facility	No.
Health post/center/clinic	19
Public hospital	9
Christian Health Association of Malawi hospital/ health center	7
Integrated health center	6
Total number of facilities	41

health posts, centers, or clinics, which provided primary health care, and public hospitals, which offered in- and outpatient services at the local, district, and regional levels. Six additional public facilities were designated "integrated" health centers and were receiving targeted support from the United Nations Population Fund (UNFPA) to fully integrate primary health care. UNFPA funded the implementation of services in these facilities, including ensuring adequate stocks of drugs and providing in-service training. Seven private facilities were part of the Christian Health Association of Malawi (CHAM), which is a network of church-owned health facilities and hospitals, and the largest private provider of health care in Malawi. Two of the 7 CHAM facilities were Catholic and did not provide modern contraceptives.

The health centers/posts tended to be lower-volume sites staffed by 1 or 2 providers offering a range of primary health services (fully integrated), whereas the rural or urban hospitals tended to be higher-volume sites where specific services were provided by different health care workers. There was a great deal of variety of how family planning and HIV services were organized among all the facilities, and in many cases, facilities were using more than 1 model of family planning-HIV service integration. The government of Malawi had several policies, guidelines, and strategies that addressed integration of family planning into HIV services, but did not explicitly dictate models of integration;[12] hence, a variety of models were being implemented.

Facilities in Malawi varied greatly in how they organized family planning and HIV services.

The approaches to service delivery integration included models where: (1) a client received multiple services from 1 provider in 1 room (often observed at UNFPA-supported integrated health centers), (2) the client might be seen by more than 1 provider in the same room or clinic space (often observed at health centers), (3) the client could access HIV and family planning services in different spaces (clinics) by different health care providers, but on the same day on the same facility grounds (often observed at public hospitals), or (4) completely non-integrated services where the client could only access HIV and family planning services on different days at that facility (often noted at health posts/clinics), or (5) needed to be referred elsewhere for their desired family planning services (in the case of CHAM facilities, or where clients wanted a family planning method not available at that site). Many facilities did not provide a range of family planning services/methods. Instead, Banja La Mtosogolo (a Marie Stopes International affiliate) would provide outreach services at these facilities at intermittent intervals (e.g., once per month) for clients interested in long-acting and permanent methods.

Study Design

Data Collection

Both quantitative and qualitative data were collected between April and May 2015 using the following methods.

Facility Audits. At the 41 facilities, data collectors used a checklist to note the facility structure, observe patient flow in the ART clinic, observe the counseling and treatment spaces in the absence of clients, and note informational materials available at the site on the day of the data collection visit. The audits included a review of basic supplies/commodities at the ART clinic and a review of the ART registers where information on the clients who had been to the ART clinic that day were being registered (N=41).

Staff Interviews. Data collectors interviewed the facility in-charge (N=41) and up to 3 health care providers responsible for providing HIV services (N=122). The facility in-charges included doctors, registered nurses/midwives, clinical officers, and paramedical workers. The health care providers responsible for providing HIV services were nurses, midwives, clinical officers, health surveillance assistants, and paramedical workers, which included auxiliary nurses, medical assistants, or a nurse-midwife technician. On the day of the interview, the providers serving clients with HIV were first approached and interviewed. These providers cater to a range of HIV services including dealing with communicable diseases/tuberculosis, pediatric patients with HIV, and HIV counseling and testing (upon receiving referrals). A maximum of 3 providers per facility were interviewed.

Client Exit Interviews. Clients attending the ART clinic were invited to participate in the study

after their regular visit for the day. Women ages 18–49 and men ages 18–59 who could read and write met the inclusion criteria. Literacy was an inclusion criteria to be able to sign the consent form. Even though the clients were randomly selected, we oversampled the number of women (compared to men) interviewed at each facility to better understand the needs and patterns of contraceptive use among patients with HIV. Based on the national prevalence of unmet need for family planning of 26% (DHS, 2010), we note that a sample size of 400 clients gives us statistical power with 95% confidence level and 5% margin of error to measure differences in family planning use among ART clients. Data collectors administered an exit interview to 425 clients.

Mystery Clients. To obtain a better understanding of client-provider interactions and referral mechanisms, 9 mystery clients (3 per region: 2 female and 1 male) were deployed to 20 facilities on days the data collection team was not visiting. These clients presented themselves as transfer patients with HIV seeking antiretrovirals (ARVs) and were trained to document whether they were spontaneously counseled and offered family planning (received PIFP as per the Government of Malawi clinical guidelines). If the health care provider did not offer family planning (i.e., if there was no provider initiation of family planning during ART services), the mystery clients were then to ask about family planning and document the provider's response. Likewise, if the provider offered condoms and/or injectables, the mystery clients were also trained to ask about another family planning method and to document how the provider responded (e.g., did the provider offer a referral). Nine individuals made 58 mystery client visits to 20 facilities across all 3 regions. The mystery clients presented themselves as ART clients temporarily in the area and in need of ARV resupply (e.g., visiting a sick relative, husband just transferred). The female mystery clients were 20–36 years old. The male mystery clients were 19, 33, and 35 years old.

The mystery clients were trained to first see whether providers mentioned family planning (PIFP), and if not, to ask about it. They were provided with suggestions for different profiles or scenarios regarding their reproductive intentions. For example, the older women said they had 3 or 4 children and didn't want any more, whereas younger women were told to say they had 1 child and wanted to space their births. The 19-year-old man presented himself as a student. The research team (with knowledge of the Ministry of Health) created temporary health passports for the clients to support their profile. Any ARVs collected by the clients were documented and returned to the health system via the Lighthouse Clinic in Lilongwe.

Data Entry, Cleaning, and Analysis

Quantitative data from facilities were collected using paper data collection forms, then entered into templates developed in CSPro6.1 and exported into STATA10 for analysis. Qualitative data were transcribed and then translated into English.

Ethical Considerations

The study received ethical approval from Malawi's National Health Sciences Research Committee in Lilongwe, Malawi, and the Institutional Review Board of Health Media Lab in Washington, DC, USA. Interviews were conducted in a private space and lasted under 1 hour. All participants (facility in-charges, providers, and clients) were provided details of the study in advance, and read aloud the consent form, which they then signed. No names were recorded, only titles, or in the case of clients, basic sociodemographic data. Providers and clients at the facilities were not given any compensation for their participation in this study. All informed consent information and subsequent questionnaires were translated and administered in one of the prevalent local languages of the region: Chichewa, Chitumbuka, or Yao.

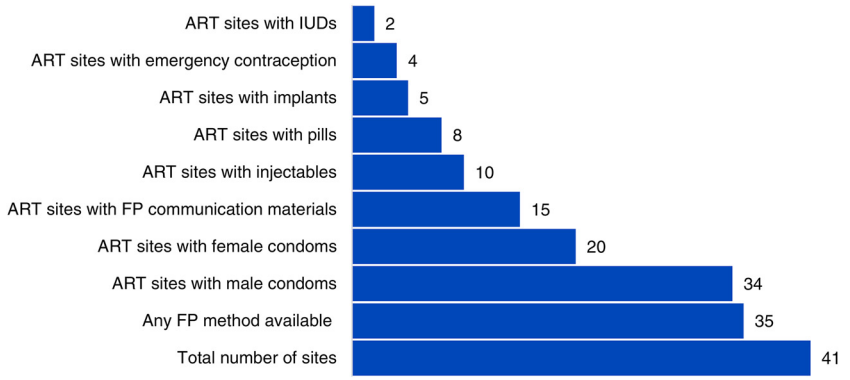
RESULTS

Availability of Family Planning Information and Commodities

Of the 41 facilities with an ART clinic or an outpatient department where ART services were being provided, 85% (n=35) had family planning methods available, but this was mainly because most facilities had condoms available, which counts as a family planning method even if primarily used for prevention of sexually transmitted infections, including HIV. Very few facilities had a variety of family planning methods available. Of the 35 facilities where family planning was available, 10 of those sites (24%) had injectables and condoms. Only 5 sites offered implants and only 2 offered intrauterine devices (IUDs). Six of the facilities (15%) had a range of hormonal, non-hormonal, and short- and long-acting contraceptives available. Furthermore, only 15 (37%) of ART clinics had family planning information materials displayed (Figure 1).

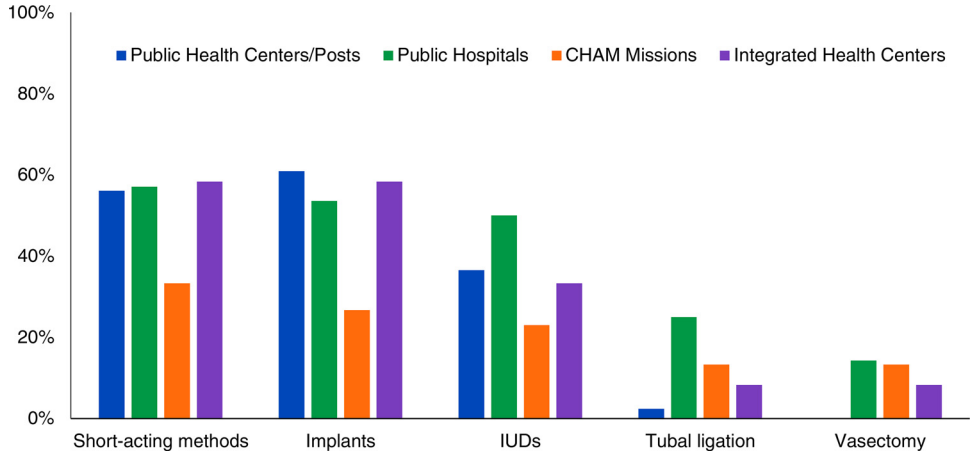
Of the 35 facilities where family planning was available, 6 had a range of hormonal, non-hormonal, and short- and long-acting contraceptives available.

FIGURE 1. Availability of Family Planning Methods in 41 Facilities in 9 Districts in Malawi, April–May 2015



Abbreviations: ART, antiretroviral therapy; FP, family planning; IUD, intrauterine device.

FIGURE 2. Percentage of Providers in 9 Districts in Malawi Who Reported Receiving Family Planning Training on Specific Contraceptive Methods, by Method and Facility Type, April–May 2015 (N=96)



Abbreviations: CHAM, Christian Health Association of Malawi; IUD, intrauterine device.

Organization of Services

Interviews with 122 health care providers found that 93% had received training on HIV services, but only 79% had received any training in family planning, and only 24% of the health providers had received any training specifically on integration of family planning and HIV services. In addition, fewer providers working in CHAM facilities had received family planning training (only 71% had been trained in family planning compared to 76%–88% of the providers in other facilities), and of those who received family planning training, fewer had received training in short-acting methods, implants, or IUDs compared to the providers at the other facilities. Overall, 53% of the providers with some pre- or in-service family planning training had been trained on short-acting methods, 38% on IUDs, and 21% of those in cadres eligible to provide sterilization had been trained on permanent family planning methods (Figure 2).

When health care providers were asked if the ART services had received some reorganization to accommodate family planning services, 83% of providers said yes (Table 2). When asked about different ways services had been reorganized,

TABLE 2. Description of Organization of Antiretroviral Therapy and Family Planning Services, According to the Health Service Provider Within Selected Facilities, by Facility Type, Malawi

	Health Centers/Posts No. (%)	Public Hospitals No. (%)	CHAM Missions No. (%)	Integrated Health Centers No. (%)	Total No. (%)
Total number of providers	54 (44.3%)	32 (26.2%)	21 (17.2%)	15 (12.3%)	122 (100%)
Have ART services been reorganized to accommodate FP services?					
Yes	43 (79.6%)	26 (81.3%)	17 (80.9%)	15 (100%)	101 (82.8%)
No	11 (20.4%)	6 (18.8%)	4 (19.0%)	0	21 (17.2%)
How have ART services been reorganized to accommodate provision of FP services?[a]					
Created more space	8 (18.6%)	4 (15.4%)	3 (17.7%)	4 (26.7%)	19 (18.8%)
Revised ART on-site protocols to accommodate FP services	12 (27.9%)	17 (65.4%)	5 (29.4%)	8 (53.3%)	42 (41.6%)
Trained ART providers in different FP methods	21 (48.8%)	11 (42.3%)	4 (23.5%)	12 (80.0%)	48 (47.5%)
Created informal referral agreements within facility	26 (60.5%)	14 (53.8%)	6 (35.3%)	5 (33.3%)	51 (50.5%)
Developed facility referral agreements across facilities	14 (32.6%)	10 (38.5%)	5 (29.4%)	2 (13.3%)	31 (30.7%)
Revised ART client registers to accommodate FP services	4 (9.3%)	4 (15.4%)	1 (5.9%)	2 (13.3%)	11 (10.9%)
Adjusted operating time for ART services	7 (16.3%)	2 (7.7%)	2 (11.8%)	4 (26.7%)	15 (14.9%)
Provided ART/FP on same day	3 (6.9%)	3 (11.5%)	1 (5.9%)	0	7 (6.9%)
Do you have time/opportunity to counsel ART clients on FP methods?					
Yes	51 (94.4%)	29 (90.6%)	20 (95.5%)	14 (93.3%)	144 (93.4%)
No	2 (3.7%)	2 (6.3%)	1 (4.8%)	1 (6.7%)	6 (4.9%)
Not sure	1 (1.9%)	1 (3.1%)	0	0	2 (1.6%)
What FP methods do you counsel ART clients on?[a,b]					
Fertility awareness methods	51 (100%)	29 (100%)	20 (100%)	14 (100%)	144 (100%)
Pills	44 (86.3%)	21 (72.4%)	18 (90.0%)	12 (85.7%)	95 (83.3%)
Male condoms	51 (100%)	26 (89.7%)	20 (100%)	14 (100%)	111 (97.4%)
Female condoms	45 (88.2%)	25 (86.2%)	19 (95.0%)	13 (92.9%)	102 (89.5%)
Injectables	47 (92.2%)	26 (89.7%)	20 (100.0%)	14 (100%)	107 (93.9%)
IUD	28 (54.9%)	18 (62.1%)	12 (60.0%)	5 (35.7%)	63 (55.3%)
Implants	43 (84.3%)	19 (65.5%)	16 (80.0%)	10 (71.4%)	88 (77.2%)
Female sterilization	34 (66.7%)	20 (69.0%)	11 (55.0%)	7 (50.0%)	72 (63.2%)
Male sterilization	25 (49.0%)	13 (44.8%)	7 (35.0%)	5 (35.7%)	50 (43.9%)
Emergency contraception	19 (37.3%)	17 (58.6%)	10 (50.0%)	5 (35.7%)	51 (44.7%)
Are clients referred out for services?					
Yes	48 (88.9%)	15 (46.9%)	16 (76.2%)	12 (80.0%)	91 (74.6%)
No	6 (11.1%)	17 (53.1%)	5 (23.8%)	3 (20.0%)	31 (25.4%)
What prior knowledge do you have of facilities to which you are referring clients for FP services?[a,c]					
Services provided	32 (66.7%)	10 (66.7%)	12 (75.0%)	9 (75.0%)	63 (69.2%)
Weekdays on which services are provided	34 (70.8%)	11 (73.3%)	6 (37.5%)	7 (58.3%)	58 (63.7%)

Continued

TABLE 2. Continued

	Health Centers/Posts No. (%)	Public Hospitals No. (%)	CHAM Missions No. (%)	Integrated Health Centers No. (%)	Total No. (%)
Times when services are provided	25 (52.1%)	5 (33.3%)	5 (31.3%)	5 (41.7%)	40 (44.0%)
Transport costs to reach the referral site	14 (29.2%)	3 (20.0%)	5 (31.3%)	4 (33.3%)	26 (28.6%)
No prior knowledge	5 (10.4%)	5 (33.3%)	1 (6.3%)	2 (16.7%)	13 (14.3%)

[a] Categories are not mutually exclusive.
[b] Of the 114 providers who counsel ART clients on FP.
[c] Of the 91 providers who refer clients for services.
Abbreviations: ART, antiretroviral therapy; CHAM, Christian Health Association of Malawi; FP, family planning.

42% of providers said that ART protocols had been revised. On-site protocols are copies of operational guidelines for each type of facility within which the HIV clinical guidelines are incorporated. As well, 51% of providers said that they had established informal referral agreements between their ART clinic and those providing family planning services either within the facility or at a public facility nearby providing family planning services. Only 15% of providers mentioned that ART provision time was adjusted to accommodate family planning, and 11% of providers reported that the ART registers had been revised.

At the time of data collection, ART registers were supposed to already have 3 columns to indicate whether family planning counseling, condoms, and/or injectables were provided. These columns were added after the clinical guidelines were first introduced. Further, this was mandated by the HIV & AIDS Department of the Ministry of Health as injectables and condoms were being dispensed through the ART essential drug list. This study investigated whether there were any other columns added to the ART registers corresponding to additional family planning methods to determine whether any additional methods beyond condoms or injectables were provided to clients on ART. When data collectors requested to view the registers at ART clinics, half (17) were unavailable—either providers would not allow data collectors to review, the register was not yet out of the locked cabinet for the day (despite patients being seen), there was a shortage of registers at the clinic, or it was at another location (or lost/misplaced). Of the ART registers reviewed at 18 sites, 6 had extra columns added in the ART register to document family planning provision, whereas another 8 maintained a separate family planning register. The remaining 4 facilities had no mechanism to document additional family planning service provision (beyond condoms or injectables) at the ART clinic. There are no other documents that capture information about clients coming to the facility to receive ART medication. Malawi's Department of HIV & AIDS regularly monitors the number of condoms and injectables distributed to ART clients.

Nonetheless, 93% of providers reported they had the time and opportunity to counsel ART clients on family planning methods available to them (Table 2), and almost all mentioned condoms and injectables as the methods they counseled on. Supplement 2 details precise wording for questions. Fewer providers mentioned counseling on implants (77%), IUDs (55%), female sterilization (63%), or vasectomy (44%) (Table 2). About 75% of providers mentioned they referred clients for family planning to other facilities, but most lacked knowledge of details about when those services were available or the transport costs to those services (Table 2).

Clients' Reproductive Intentions and Contraceptive Needs

We conducted 425 exit interviews with HIV clients (at least 10 per site across all 41 sites). Of these clients, 419 of them disclosed they were HIV-positive: 332 were female, and 87 were male. Of the female clients interviewed (n=332), 17 (5%) were currently pregnant, and most reported the pregnancy as either mistimed (n=9) or unwanted (n=4). Of the 315 female clients who were not pregnant, 52% reported not wanting any more children, 14% wanted to wait more than 2 years to get pregnant, and an additional 25% were unsure. Fifty clients (16%) reported already using sterilization as their family planning method.

Of the total number of male and female clients eligible to use family planning (i.e., themselves or partner not pregnant, not already sterilized, n=358), 60% were using a method to avoid pregnancy (this is comparable to the modern contraceptive prevalence of married women in Malawi reported in the 2015/16 DHS of 58%). Half were using male condoms, and one-third were using injectables. Only 11% were using implants, about 4% were using female condoms, 4% were using pills, and only 1 client (0.5%) was using an IUD.

When asked about their preferences for receiving integrated services, 97% of clients interviewed preferred receiving fully integrated services (i.e., in the same clinic/same room, same day). In addition, 90% of clients said that they would be willing to wait longer to get multiple services per visit. Over three-quarters of clients stated making fewer trips to the facility as the benefit of receiving integrated family planning-HIV services, and 43% cited reduced travel costs as a benefit of integration. Less than 10% of clients mentioned reduced stigma as a benefit of integrating services, even though theoretically, integrated services can reduce HIV-related stigma because other clients and providers would not explicitly know the nature of the service the client was seeking.

Provider-Initiated Family Planning

To ascertain whether PIFP was being implemented in ART clinics as stipulated in the national clinical guidelines, we specifically asked clients who had come for ART (n=355) and other HIV services (n=51) if any health provider had asked them about their fertility intentions and/or offered them family planning during their visit that day (Supplement 2, Question 2). Only 56 (14%) said yes.

In another line of questioning, we asked clients if a provider had *ever* inquired about their fertility intentions (Supplement 2, Question 3). Of the 77 clients that said yes, 13 reported "every time," 21 said "often," 11 said "sometimes," and 23 said "rarely."

Mystery Client Visits

Of 58 visits, only 2 of the mystery client visits reported the provider proactively offered family planning to the clients:

> *She [the provider] noted my book had nothing on family planning and started advising me of family planning and all methods like vasectomy, Norplant [outdated term for implants], IUCD [intrauterine contraceptive device] . . . she later advised me to opt for a family planning method to avoid unwanted pregnancy.* — Female, 20, health center

> *She [the provider] said "All of [the choices] are present," and added it was my choice to choose which one I prefer.* — Female, 36, district hospital

Eleven mystery client visits at 7 different facilities resulted in the mystery clients being told that they could not receive these ART services since they were not registered as regular clients at the facility. For the most part, the mystery clients did not encounter a conducive or welcoming environment for PIFP. The young male mystery client reported not being taken seriously at 2 facilities when he asked about family planning and was only offered condoms:

> *I then asked for family planning to which he [the provider] responded how come I wanted family planning when I was in school [provider offered family planning options and information, but laughed at him].* — Male, 19, health center

Clients cited fewer trips to facility, reduced travel costs, and reduced stigma as benefits of integrated family planning and HIV services.

Another health center fared particularly poorly in their interaction with the mystery clients. The 2 quotes below are from the same location:

> *When I asked him [the provider] about family planning he shouted at me saying the room was not for family planning:* "*Had it been that you are looking for family planning you could have gone to the family planning room. Go out, I want to assist other patients please.*" *I ask him about condoms. He said I am wasting his time, there was no condoms.* — Male, 33, health center

> *Then I asked about family planning and I was told that I should not delay him [the provider] as he has a lot of work to do and he sent me away. He said that if I want family planning methods I should come the following day around 8 a.m.* — Female, 24, health center

Only 14% of clients who accessed HIV services said that providers had asked them about fertility and offered family planning.

DISCUSSION

This study used a mixed-methods approach to triangulate and validate information from providers, clients, and facility audits in an effort to reveal whether PIFP in ART clinics was truly being implemented as envisioned by the 2011 (and 2014) clinical management guidelines in Malawi. For the most part, this service delivery approach has largely been unrealized in practice.

Although 93% of providers reported having enough time for counseling ART clients on family planning, only 14% of clients reported being asked about their fertility intentions or being counseled

Although 93% of providers reported having time to counsel clients on family planning, only 14% of clients reported being asked about their fertility intentions or being counseled about family planning.

on family planning at the ART clinic that day. Using mystery clients is a valuable approach to obtaining information on client-provider interactions.[13,14] It allows researchers to test how services are provided given certain client profiles, minimizes recall or other biases in self-reporting through interviews, and reduces the "Hawthorne Effect"—that data collectors undertaking observational assessments may influence provider and client interactions merely by their presence. Therefore, this study also conducted mystery client visits in a subset of the facilities. Only 2 of 58 (3%) mystery client visits resulted in PIFP being offered to the client.

The client exit interviews revealed a significant need for family planning services among clients with HIV. Our findings that the vast majority of female clients wanted no more children or wanted to delay childbearing for more than 2 years echo multiple other studies on the reproductive intentions of women with HIV.[6–8] In particular, we found that 13 of 17 of the pregnant women reported the pregnancy as mistimed or unwanted. Although this was a small sample size, it echoes a recent article from Malawi on pregnant women on ART (N=220) that found 75% of women reported the pregnancy as mistimed or unwanted.[9] Through client exit interviews, clients with HIV expressed a significant interest in receiving integrated services. Almost all said they would be willing to wait longer to receive multiple services to reduce trips to the facility and transportation costs. This finding suggests the financial and opportunity costs of seeking health care may be more onerous for Malawians than managers of the health system realize.

The demand for integrated family planning-HIV services and the high unmet need for family planning among women with HIV underscores the need for more focused efforts to implement provider-initiated family planning in Malawi's HIV services.

The demand for integrated family planning-HIV services and the high unmet need for family planning among women with HIV in Malawi underscores the need for more focused efforts to implement PIFP in Malawi's HIV services. This will require more concrete programmatic interventions to strengthen and sustain family planning-HIV integration, as well as further investigation into the challenges providers and facility managers face in institutionalizing PIFP, such as time constraints and lack of training, inadequate organization of services, stock-outs, and insufficient accountability for implementing national guidelines. Exploring provider perspectives on family planning-HIV integration and possible biases or negative attitudes about PIFP would also be helpful to understand what additional interventions or trainings are needed. Although 93% of providers said they had enough time to discuss their clients' family planning needs (Supplement 2, Question 1) our triangulation of data from exit interviews, and mystery clients suggest that clients only see providers for less than 5 minutes on average. Thus, the authors conclude that the providers' response of adequate time was likely due to a social desirability bias. It is also possible that providers at ART clinics are not motivated to provide PIFP. Facilities/providers currently implementing PIFP focus largely on male and female condoms and injectables, which certainly is the emphasis of the national clinical guidelines. Furthermore, it is not surprising that providers focus on condoms and injectables because providers are likely conditioned to promote condoms (for HIV prevention), and in addition to being part of the essential ART drug list, injectables are the most common family planning method in Malawi. Providers need to be able to counsel clients with HIV on a full range of family planning methods, provide some method choice, and effectively refer as needed. This study found that although 77% of the providers reported counseling on pills or implants, fewer mentioned female sterilization (63%), IUDs (55%), or vasectomy (44%). This study also found only about a third of facilities had injectables available where ART services were provided, and family planning communications materials were also scarce. This lack of method availability affects informed choice. However, HIV clients, like everyone else, deserve comprehensive family planning counseling and services and access to a full range of methods to meet their reproductive intentions and life context, which likely change over time.

Limitations

We note the following limitations of this study. Response bias may have affected our data collected from in-person interviews. Facility in-charges and providers may have overstated the level of family planning-HIV integration of services and availability of time for counseling or underestimated stock-outs. Clients in exit interviews may have been subject to recall bias. Furthermore, the presence of data collectors at the facility conducting interviews and conducting the facility audit may also have produced a Hawthorne effect with providers and clients, affecting service delivery approaches, though no direct observation of client-provider interactions occurred. Our use of mystery clients encountered some challenges we did not anticipate. The mystery clients faced challenges in being served at facilities due to presenting as ad hoc or "emergency" clients, and a few were refused services outright.

CONCLUSION

Although Malawi should be recognized as an early adopter of PIFP within its HIV-management guidelines, 4 years after adopting these guidelines, implementation of PIFP was largely unrealized at the clinic level. The results of the larger study on which this article is based were shared with the Ministry of Health and USAID in-country at a national conference and to a wider audience of implementers and researchers in 2 global meetings and conferences. Informal follow-up with in-country contacts during the drafting of this article suggests that since the collection of data reported in this study, the Government of Malawi has conflicting information on integrating family planning into HIV at the policy level. For example, in 2016, Malawi issued a third edition of its *Clinical Management of HIV in Children and Adults,* which maintained PIFP as a protocol during ART services. However, the 2016 update to the ART registers removed columns to report condom or injectable provision. As such, it is likely that this article's findings on PIFP at the facility level remain applicable and further institutionalizing of PIFP in Malawi's public and private health facilities requires targeted and comprehensive systems changes. Only about a quarter of providers said that they had been trained on family planning-HIV integration, and only half had information on where to refer clients for their family planning needs. These findings suggest that providers may not be fully aware of how to implement PIFP; hence further training may be warranted. Simple job aids to reinforce PIFP and support referrals may also be helpful. Routine follow-up of patients during subsequent visits by their HIV service providers to see if they are having their family planning needs met, and a more robust provision of a full range of family planning methods accessible to each client either within the ART clinic or through a referral, would also help support HIV clients to meet their reproductive intentions. In addition, facility in-charges and health management teams should be held accountable to measurable PIFP indicators. This can be done through the routine review of data among hospital staff and open dialogue on how to make family planning services more available to ART clients. Specifically, there is a need for improved tracking and reporting of family planning commodities and services provided through ART services, expanded family planning method choice where possible, and strengthened referral systems. Continuous availability of family planning commodities at ART clinics coupled with a formalized referral system will allow the ART services to be more integrated and hence meet the needs of ART clients. Finally, community sensitization and demand-creation for family planning-HIV integrated services might help inform clients with HIV about their reproductive health rights and empower them to proactively ask for family planning during their visits to the clinic; this may also help to promote more accountable health facilities where PIFP is not being practiced.

Acknowledgments: The authors would like to acknowledge Olive Mtema and Pierre Dindi for overseeing the original study; Nyanyiwe Mbeye and Fannie Kachale for assisting with the study design and institutional review board application; Betty Farrell for review of the data collection tools; all the data collectors and mystery clients involved in the study, particularly our data quality assurance consultants, Julius Chingwalu, Esther Kip, and Albert Dube; and the district health officers and all the respondents who gave their time to participate in the study. We would like to acknowledge the valuable technical review we received from our colleagues Sara Bowsky, Jay Gribble, and Ron MacInnis in design and write-up of the study. We appreciate the advice, encouragement, and assistance of our colleagues Jay Gribble, Olive Mtema, and John Stanback during the development of this article.

Competing interests: None declared.

REFERENCES

1. EngenderHealth. *Integrating Family Planning and Antiretroviral Therapy: A Client-oriented Service Model.* New York: EngenderHealth; 2014. https://www.engenderhealth.org/files/pubs/family-planning/integrating-family-planning-antiretroviral-therapy-service-model.pdf. Accessed March 23, 2019.
2. The Challenge Initiative. Provider-initiated Family Planning (PIFP). https://tciurbanhealth.org/wp-content/uploads/2017/06/Provider-Initated-FP.pdf. Accessed March 23, 2019.
3. Meeting the sexual and reproductive health needs of women living with HIV. What Works for Women website. http://www.whatworksforwomen.org/chapters/15-Meeting-the-Sexual-and-Reproductive-Health-Needs-of-Women-Living-With-HIV/sections/35-Meeting-the-Sexual-and-Reproductive-Health-Needs-of-Women-Living-With-HIV/evidence#s-223. Accessed March 17, 2019.
4. World Health Organization (WHO) and Joint United Nations Programme on AIDS (UNAIDS). 2007. *Guidance on Provider-Initiated HIV Testing and Counseling in Health Facilities.* Geneva: WHO. http://apps.who.int/iris/bitstream/handle/10665/43688/9789241595568_eng.pdf;jsessionid=D28BDF0AA411D1F98CACFCAD83380BDE?sequence=1. Accessed March 17, 2019.
5. Kennedy CE, Fonner VA, Sweat MD, Okero FA, Baggaley R, O'Reilly KR. Provider-initiated HIV testing and counseling in low- and middle-income countries: a systematic review. *AIDS Behav.* 2013; 17(5):1571–1590. CrossRef. Medline
6. Clayden P. High rate of unintended pregnancy among HIV positive African women on ART. Conference Report of paper presented at: AIDS 2018; July 23–27, 2018; Amsterdam, The Netherlands. http://i-base.info/htb/34980. Accessed March 17, 2019.
7. MacQuarrie KLD. *HIV/AIDS and Unmet Need for Family Planning.* Rockville, MD: ICF International; 2015. https://www.dhsprogram.com/pubs/pdf/WP122/WP122.pdf. Accessed March 17, 2019.
8. Wanyenze RK, Matovu JKB, Kamya MR, Tumwesigye NM, Nannyonga M, Wagner GJ. Fertility desires and unmet need for family planning among HIV-infected individuals in two HIV

clinics with differing models of family planning service delivery. *BMC Women's Health*. 2015;15:5. CrossRef. Medline

9. O'Shea MS, Rosenberg NE, Tang JH, et al. Reproductive intentions and family planning practices of pregnant HIV-infected Malawian women on antiretroviral therapy. *AIDS Care*. 2016;28(8):1027–1034. CrossRef. Medline
10. Malawi Ministry of Health. *Clinical Management of HIV in Children and Adults*. Lilongwe, Malawi: Malawi Ministry of Health; 2014. https://aidsfree.usaid.gov/sites/default/files/tx_malawi_2014.pdf. Accessed July 31, 2019.
11. Irani L, McGinn E, Melish M, Mtema O, Dindi P. *Integration of Family Planning and HIV Services in Malawi: An Assessment at the Facility Level*. Washington, DC: Futures Group, Health Policy Project; 2015. https://www.healthpolicyproject.com/pubs/3876_FINALMalawiHIVFacilityReport.pdf. Accessed March 17, 2019.
12. Irani L, Pappa S, Dindi P. *Family Planning and HIV Integration in Malawi: A Policy Analysis*. Washington, DC: Health Policy Project. http://www.healthpolicyproject.com/pubs/455_FINALPolicyReviewFPHIVIntegration.pdf. Accessed March 17, 2019.
13. Boyce C, Neale P. *Using Mystery Clients: A Guide to Using Mystery Clients for Evaluation Input*. Watertown, MA: Pathfinder International; 2006. http://www2.pathfinder.org/site/DocServer/m_e_tool_series_mystery_clients.pdf. Accessed March 23, 2019.
14. Fitzpatrick A, Tumlinson K. Strategies for optimal implementation of simulated clients for measuring quality of care in low- and middle-income countries. *Glob Health Sci Pract*. 2017;5(1):108–114. CrossRef. Medline

Peer Reviewed

Received: May 23, 2019; **Accepted:** September 16, 2019; **First Published Online:** November 27, 2019

Cite this article as: McGinn EK, Irani L. Provider-initiated family planning within HIV services in Malawi: did policy make it into practice? *Glob Health Sci Pract*. 2019;7(4):540-550. https://doi.org/10.9745/GHSP-D-19-00192

GLOBAL HEALTH: SCIENCE AND PRACTICE
Dedicated to what works in global health programs
OPEN ACCESS

ORIGINAL ARTICLE

Realizing the "40 by 2022" Commitment From the United Nations High-Level Meeting on the Fight to End Tuberculosis: What Will It Take to Meet Rapid Diagnostic Testing Needs?

Amy S. Piatek,[a] William A. Wells,[a] Kaiser C. Shen,[a] Charlotte E. Colvin[a]

Existing rapid diagnostics offer faster and more sensitive diagnosis of tuberculosis (TB) and simultaneous detection of multidrug-resistant TB. A 5-fold increase in investment in these tools is needed to meet the needs of the TB community and the United Nations' ambitious 40 million by 2022 diagnosis and treatment target.

ABSTRACT

The potential gains from full adoption of World Health Organization (WHO)-recommended rapid diagnostics (WRDs) for tuberculosis (TB) are significant, but there is no current analysis of the additional investment needed to reach this goal. We sought to estimate the necessary investment in instruments, tests, and money, using Xpert MTB/RIF (Xpert), which detects *Mycobacterium tuberculosis* (MTB) and tests for resistance to rifampicin (RIF), as an example. An existing calculator for TB diagnostic needs was adapted to estimate the Xpert needs for a group of 24 countries with high TB burdens. This analysis assumed that countries will achieve the case-finding commitments agreed to at the recent United Nations High-Level Meeting on the Fight to End Tuberculosis, and that countries would adopt the WHO-recommended algorithm in which all people with signs and symptoms of TB receive an Xpert test. When compared to the current investments in these countries, this baseline model revealed that countries would require a 4-fold increase in the number of Xpert modules and a 6-fold increase in the number of Xpert test cartridges per year to meet their full testing needs. The incremental cost of the additional instruments for these countries would total approximately US$474 million, plus an incremental cost each year of cartridges of approximately $586 million, or a 5-fold increase over current investments. A sensitivity analysis revealed a variety of possible changes under alternative scenarios, but most of these changes either do not meet the global goals, are unrealistic, or would result in even greater investment needs. These findings suggest that a major investment is needed in WRD capacity to implement the recommended diagnostic algorithm for TB and reach the case-finding commitments by 2022.

INTRODUCTION

We are at a time of unprecedented attention and opportunity for tuberculosis (TB). At the United Nations High-Level Meeting on the Fight To End Tuberculosis on September 26, 2018,[1] world leaders committed to bold targets and urgent action to end TB, including diagnosing and treating a cumulative 40 million people by 2022 (40 by 2022).[2] This commitment cannot be met without equally bold and urgent responses to the greatest challenges facing national TB programs including the continued lack of access to quality and rapid TB diagnostics for people with signs and symptoms of TB. In 2017, 3.6 million people with TB went either undiagnosed or were detected and not reported, representing 36% of the estimated 10 million new cases.[3]

Rapid and accurate diagnosis is a critical requirement for an effective TB care and prevention effort[4] because delayed diagnosis results in greater morbidity and mortality and increased disease transmission. For too long, TB programs had access only to smear microscopy, a century-old technology that has low sensitivity, detecting only about half of all TB cases (fewer in paucibacillary disease) and not detecting drug-resistant TB at all.[5,6] Although we now have rapid TB diagnostic tests, the majority of people in high TB burden countries continue to be tested for TB with smear microscopy.[3]

Multiple documents from the global TB community have outlined the target of universal access to World Health Organization (WHO)-recommended rapid diagnostics (WRDs). Pillar 1 of the End TB Strategy[7] states that early diagnosis of all persons of all ages with any form of TB is fundamental and that WRDs and drug

[a] United States Agency for International Development, Washington, DC, USA. Correspondence to William Wells (wwells@usaid.gov).

Multiple documents from the global TB community have outlined the target of universal access to WHO-recommended rapid diagnostics.

susceptibility testing should be available to all who need it. WHO recommends that TB programs use WRDs that allow for the simultaneous detection of TB and drug-resistant TB as the initial diagnostic test instead of smear microscopy (e.g., in the compendium of WHO guidelines and standards[8]; see methods section for further details). Finally, the Global Laboratory Initiative's Model TB Diagnostic Algorithms,[9] originally published in early 2017 and recently revised, provides the preferred algorithm for universal patient access to rapid testing to detect *Mycobacterium tuberculosis* (MTB) and resistance to the anti-TB drug rifampicin (RIF). The algorithm currently indicates that the initial diagnostic test to use is the Xpert MTB/RIF (Xpert) assay, including for use with adults and children with signs and symptoms or chest X-ray (CXR) suggestive of TB, with persons being evaluated for extrapulmonary TB, and with persons being evaluated for TB who are HIV-positive.

Initial roll-out of Xpert began in the second quarter of 2010, and by the end of that year, 18 countries had 166 modules in place and cumulatively had run approximately 27,000 tests (Philippe Jacon, Cepheid, email communication, Novem-ber 2018). By the end of 2017, 130 countries had procured 'more than 42,000 modules and 34.4 million tests.[3] However, in countries where data were available and reported, only 20.6%[10] of new and relapse cases were tested using WRDs in 2017.

Effective TB case finding first requires a variety of approaches to identify all people with signs and symptoms of TB.[11–13] Although this initial identification of presumptive TB is critically important to reaching the 40 by 2022 targets, the focus of the current study is to investigate whether or not TB diagnostic networks in high-burden countries have the capacity to provide rapid and accurate testing for an initial TB diagnosis for these clients and to determine the actual requirements to test all people with signs and symptoms of TB with a WRD. This effort requires a realistic calculation of what volumes of testing and numbers of testing instruments a country needs. We present such an analysis using Xpert as an example because it is the most widely implemented WRD currently (although the lessons should apply equally to other WRDs).[14]

METHODS

We adapted the WHO TB diagnostic capacity calculator[15] (WHO calculator) to generate estimates of the need for rapid TB diagnostic testing instruments and assays using the Xpert instrument and Xpert MTB/RIF cartridges as an example. We estimated the volume of Xpert cartridges needed to test persons presenting with signs and symptoms of TB and the number of Xpert modules that are needed to provide full coverage for TB diagnosis in a high-burden country under realistic operating conditions. The baseline model calculated the cartridges and modules needed to identify 90% of the total estimated TB cases (all forms)[3] because this reflects the level of case finding needed to reach the 40 by 2022 targets, as innovative case finding approaches are implemented at scale. The analysis was conducted for the 24 countries currently receiving direct United States Agency for International Development (USAID) funding for TB[16]; these countries represent 74% of the world's TB burden.

The WHO calculator was originally developed using a stakeholder consensus process and consists of a simple but logical Microsoft Excel format.[15] Populations who require testing are calculated from WHO epidemiological information; these figures are further multiplied based on the projected usage of Xpert. Adapting the WHO calculator methodology, which dates from a period of more conservative use of Xpert in diagnostic algorithms, the calculation was performed to determine the rapid TB diagnostic testing needs for 3 populations of TB patients: (1) HIV-negative adults; (2) HIV-positive adults; and (3) children (0–14 years old). WHO data were used to estimate the number of HIV-negative adults and the number of children with TB,[10] in both cases multiplied by 90% to reflect the 40 by 2022 targets, as noted above. The Joint United Nations Programme on HIV/AIDS (UNAIDS) 2017 data[17] were used to estimate the total number of people living with HIV (PLHIV) for each country, which was then multiplied by 81% (the number of PLHIV enrolled on antiretroviral therapy [ART] if 90% of all PLHIV know their status and 90% of all who know their status are enrolled on ART, as expected under the UNAIDS 90-90-90 treatment targets[18]). We assume that, at each of 2 visits per year, 20% of those enrolled on ART will be eligible for Xpert testing based on a symptom screen (see later justification for the 20% value).

Realistic operating capacity per module was defined as an instrument that runs 168 days per year (70% of a full working year, accounting for facility closure due to holidays or worker actions, absence of staff, or compromised power supply) and 3 tests per day (accounting for delays in specimen transport, stock-out of cartridges, staff workload, and limited operating hours).

The model, and in particular the choice of populations assumed to require Xpert, conforms to the guidance and standards found in the WHO compendium.[8] This includes Standard 6, which states that "all patients with signs and symptoms of pulmonary TB who are capable of producing sputum should have as their initial diagnostic test at least 1 sputum specimen submitted for Xpert MTB/RIF Ultra assay," and Standard 5, which confirms that "TB programmes should transition to replacing microscopy as the initial diagnostic test with WHO-recommended rapid diagnostics that allow for the simultaneous detection of TB and drug-resistant TB." Due to this assumption that all persons with signs and symptoms of TB would receive an Xpert test, the percentage of individuals that would get an Xpert test solely to test for resistance to rifampicin and not primarily for case detection was set at 0. Of note, the calculations are not expected to differ depending on use of Xpert MTB/RIF or Xpert Ultra because the algorithm would remain the same and the cost of MTB/RIF and Ultra cartridges is the same for high-burden countries. The WHO calculator's additional calculation for relapse patients was also omitted because our overall calculation was based on testing 90% of all estimated new and relapse patients, but the WHO variable "ret_nrel" (previously treated patients, excluding relapse cases) was retained. For the purposes of this analysis, the WHO calculator does not estimate any use of the Xpert instrument for other indications such as HIV viral load testing, early infant diagnosis of HIV, and/or hepatitis C testing.

Systematic reviews were conducted in both Google and PubMed for terms including ("TB" OR "tuberculosis") AND ("NNT" OR "NNS" OR "number needed to treat" OR "number needed to screen" OR "TB testing" OR "TB screening"); these results were also refined by adding search terms such as "HIV" and "X-ray." The limited relevant data resulting from this search are described in the results. Based on these findings, in the baseline model, testing 10 persons with signs and symptoms of TB with Xpert yields on average 1 diagnosed TB patient (see Results for justification). The equivalent number used for children was 4 tested with Xpert to yield 1 TB patient, based on the defaults used in the original WHO calculator.

These baseline model parameters (see Table 1) were then varied in a sensitivity analysis to cover a range of possible country-specific values and scenarios. This exercise was limited to a deterministic sensitivity analysis because the paucity of available evidence would not support the use of uncertainty distributions or of multivariate models—the latter were judged more likely to obscure rather than illuminate the critical findings.

Given the limited geographical access to Xpert testing in almost all high-burden countries, we created a separate calculation (unrelated to the WHO calculator described above) to illustrate the potential demand for Xpert based on population

TABLE 1. Baseline Model Parameters Needed to Calculate Xpert Cartridges and Modules to Identify 90% of Total Estimated TB Cases

Parameter	Value
Estimated TB burden coverage	90%
Type of WRD	Xpert MTB/RIF
Who receives a WRD?	All with TB symptoms
Number of adults with symptoms needed to test with WRD to diagnose one adult with TB	10
Number of children with symptoms needed to test with WRD to diagnose one child with TB	4
Number of days per year that WRD (module) is operational	168
Number of test cycles per day per module	3
Estimated percentage of PLHIV enrolled on ART	81%
TB screening visits per year for clients on ART	2
Percentage of clients on ART with symptoms that require WRD	20%

Abbreviations: ART, antiretroviral therapy; PLHIV, people living with HIV; TB, tuberculosis; WRD, World Health Organization-recommended rapid diagnostic.

size, regardless of estimated burden of TB and TB/HIV coinfection and operational capacity. We converted the standard for access to smear microscopy described under the Global Plan to Stop TB[19] (1 microscope per 100,000 population) to the corresponding number of Xpert sites that would be needed to provide the same level of geographic coverage. This exercise was limited to an estimation of the minimum number of Xpert sites per country and did not consider the number of Xpert modules needed per site because that would have required using variables for the number of smears per day and the reasons for those smears, and reliable data to inform such a calculation were not available.

Finally, after generating the number of modules needed in the baseline model and various scenarios, we compared these outputs to the actual numbers of modules in countries and the numbers of test cartridges procured in 2017. Because WHO does not collect these data variables as part of their annual reporting and no other standardized database exists with this information, we used alternative data sources. For the number of modules currently in countries, we surveyed national TB program staff, USAID technical representatives, and other technical partners, and compared these figures to those from other relevant country reports. For the number of test cartridges available annually in countries, we used 2017 procurement data provided by Philippe Jacon, Cepheid, manufacturer of Xpert (email communication, November 2018). The Cepheid procurement data include test cartridges procured in the public sector for 145 high-burden and developing countries[20] for MTB/RIF and MTB/RIF Ultra TB assays.

For the cost analysis, the current number of modules was subtracted from the total modules needed according to the baseline model to obtain the incremental number of modules required. This number was then divided by 4 and multiplied by the concessional cost of a 4-module machine with laptop (US$17,500 ex works, which does not include shipping or any potential customs costs; available to all of the high-TB burden countries included in this study). In line with the overall conservative approach to this analysis, variable service and maintenance costs were not included, and we used the global concessional price per test cartridge of US$9.98 ex works.

Programmatic experience supports using a ratio of at least 10 adults tested with Xpert to 1 diagnosed TB patient.

RESULTS

Determining Baseline Inputs

Below are detailed findings from the literature searches used to derive 3 of the model inputs.

Number of Adults With Symptoms Needed to Test With Xpert to Diagnose 1 Adult TB Patient

For the number of adults needed to test to diagnose 1 adult TB patient (number needed to test, or NNT), a 10:1 ratio has appeared in the guidance for many years.[21] This ratio was originally based on expert consensus with the anticipation that countries would revise based on country-specific data because it varies with epidemiology and the intensity of case finding efforts. Based on this prior use and the evidence from South Africa (see below), a ratio of 10:1 was also used in the current model. Note, however, that WHO[15,22] used a ratio of 10 tests to diagnose 1 bacteriologically positive patient (not 1 TB patient). This is a more complicated solution because the percentage of bacteriological positivity is expected to change over time with the increasing use of more sensitive diagnostics such as Xpert and Xpert Ultra, and we were not able to determine an evidence base for WHO's rationale.

A clear country example would assist in justifying this important ratio. However, evidence from many countries was found to be focused on number needed to screen (NNS, the number of individuals that need to be asked about TB symptoms to diagnose 1 TB patient) instead of the NNT figure required as an input to the baseline model. In addition, most countries are still implementing a mixture of smear microscopy and Xpert, so it becomes difficult to untangle the number needed to test with Xpert to find 1 TB patient. South Africa is more promising in this regard because it uses Xpert as the primary diagnostic test for TB.[23] Initial findings showed that the test positivity rate in South Africa jumped from the 8% seen with microscopy to 16%–18% in the first year of Xpert implementation, but this gradually declined to 12% in the fourth year,[24] and in recent years has settled on 10.2% over multiple years of measurement.[25]

These South Africa numbers represent the ratio for detecting all TB, whether in PLHIV or HIV-negative individuals. Using the U.S. President's Emergency Plan for AIDS Relief (PEPFAR)[26] and WHO[10] data from 2017 and the first half of 2018, these data can be disaggregated into an estimated 7.6% of the Xpert testing volume being used to screen PLHIV (yielding approximately 29% of the total TB cases with NNT of 2.6 due to the non-aggressive symptom screening, which also explains why only 29% of case finding was in PLHIV despite a coinfection rate of 60%), compared to an NNT of 13.7 for the remaining 61% of diagnosed TB patients. Overall, this programmatic experience

continues to support using a ratio of at least 10 adults tested with Xpert to 1 diagnosed TB patient, with the caveat that NNT varies substantially depending on prevalence (lower prevalence means higher NNT), symptom screening algorithm (a more inclusive symptom screen means higher NNT), and clinical practice.

Percentage of PLHIV on ART With Signs and Symptoms of TB Who Require Xpert

The percentage of PLHIV on ART who should be tested with Xpert varies considerably between what is ideal (and seen in study settings) versus what is typically done programmatically.

WHO recommends that PLHIV should be routinely screened for active TB at every health facility visit using a 4-symptom screen (current cough–any duration, fever, night sweats, and weight loss).[27] The presence of any 1 of these 4 symptoms is considered a positive screen, and the absence of all 4 symptoms is considered a negative screen. Thus, the current definition of optimal practice is to be broad and inclusive in the symptom screen. In countries, however, the exact symptom screen used varies, and also the patients' definition of a cough and the providers' index of suspicion vary. Thus, it is not possible to get a single, consistent number for this percentage, not just because of varying epidemiology but also because of this between-country variation.

A meta-analysis[28] includes summary statistics of 11% with cough of 2 weeks or more, 20% of PLHIV with current cough, and 47% with any 1 of current cough, fever, night sweats, or weight loss (the latter being the WHO-recommended definition of symptomatic TB among PLHIV).

In some more recent individual studies, the percentage of PLHIV or client on ART judged to have a positive symptom screen that warranted TB testing varied by country setting:

- from 5% in Ghana, in a setting with an unusually restrictive algorithm requiring cough plus 1 other symptom[29]
- to 10.5% in India, though 30% actually had at least 1 TB symptom[30]
- to 20.9% in Ghana, though only 12.6% before an intervention to increase provider awareness[31]
- to 22.9% in Rwanda, a high-screen positive percentage despite a relatively high median CD4+ of 385[32]
- to 39% in Ethiopia, despite 89% being on ART[33]
- to 53% in Kenya, screened at enrollment, including 25% with current cough[34]

Meanwhile, global PEPFAR programmatic data showed a value of only 2.8% in PEPFAR countries based on PEPFAR data from Panorama for the first half of 2018 (Sevim Ahmedov, MD, USAID, email communication, November 2018), presumably based on incomplete implementation and a low index of suspicion.

Based on this evidence, and because the model is aiming for optimal practice in terms of coverage and implementation, the baseline model includes a value for this variable of 20% (representing a value that is the median of the 3 values from the meta-analysis and is very close to the median from the 6 more recent studies). Because 20% is probably still lower than optimal, and actual implementation is closer to 3%, the sensitivity analysis also includes scenarios that cover the range of 2.5% to 30% for this variable.

Percentage of People With TB Signs and Symptoms and an Abnormal CXR Requiring Xpert

Based on WHO guidance,[35] "CXR and further clinical assessment can be used to triage who should be tested with the Xpert MTB/RIF assay to reduce the number of individuals tested and the associated costs, as well as to improve the pretest probability for TB and, thus, the predictive value of the Xpert MTB/RIF assay." Therefore, at least in theory, CXR could be used in this model to reduce the number of Xpert modules and tests needed.

We estimate that about 50% of people with signs and symptoms of TB will have any abnormality on CXR sufficiently suggestive of TB to merit further evaluation. This assumption would halve the number of Xpert cartridges and modules needed, but only if it was possible to develop CXR capacity at the subdistrict level to facilitate patient access—a massive task, and one that would increase resource needs in other ways.

The best data to support the 50% estimate would be a survey of actual CXRs from the country among people with TB symptoms. However, prevalence surveys do not typically report their data in terms of "number of people with symptoms who had an abnormal CXR." The primary input to the estimate of 50%, which was originally based on expert opinion, is from a single study that found that 45% of people with symptoms in Kenya had an abnormal CXR as read by a primary physician.[36]

Baseline Model Outputs

Table 2 presents the logic of the main calculations, and Table 3 presents the results of the baseline

TABLE 2. Calculation Logic for Baseline Model

Value to calculate	Components used in calculation	Formulae used
Total annual number of Xpert MTB/RIF tests = Number of tests for HIV+ adults, children, and HIV− adults	Number of tests for HIV+ adults	Estimated number of PLHIV[a] · 81%[b] · 2[c] · 20%[d]
	Number of tests for children	Estimated number of children with TB[e] · 90%[f] · 4[g]
	Number of tests for HIV− adults	Number of TB patients[e,h] · 90%[f] · Percentage of TB patients who are adults[e] · Percentage of adults who are HIV−[a] · 10[i]
Target number of Xpert modules needed		Total annual number of Xpert MTB/RIF tests/68[i] · 3[k]

Abbreviations: HIV+, HIV positive; HIV−, HIV negative; MTB, *Mycobacterium tuberculosis*; TB, tuberculosis; WRD, World Health Organization-recommended rapid diagnostic.
[a] Source: Joint United Nations Programme on HIV/AIDS.
[b] Target percentage of persons living with HIV who are enrolled on ART (90%· 90%).
[c] Number of TB screening visits per year for clients on ART.
[d] Percentage of clients on ART with symptoms that require testing with a WRD.
[e] Source: World Health Organization.
[f] Target for TB-burden coverage.
[g] Number of symptomatic children needed to test with WRD to diagnose 1 child with TB.
[h] All forms, including all incident, relapse, and previously treated.
[i] Number of symptomatic adults needed to test with WRD to diagnose 1 adult with TB.
[i] Number of days per year that WRD module is operational.
[k] Number of test cycles per day per module.

model. The total current number of modules in the 24 countries was 26,873 (average per country of 1,120; range 80–4,780), whereas the total modules needed in the baseline scenario was 135,198 (average per country of 5,633; range 138–49,986), suggesting the need for a 4-fold increase in the number of Xpert modules across these countries, with the percentage increase needed per country ranging from 13% (Tajikistan) to 946% (India).

Many modules are currently operating at a lower capacity than that assumed in the baseline model; therefore, the cartridge gap (the gap between the number of test cartridges procured in 2017 and the total number needed) was even greater than the module gap (the gap between the current number of modules and the number needed). A more than 6-fold increase in test cartridges would be needed to get from the current volume per year in the 24 countries of 9,404,400 (average per country of 391,850 cartridges per year; range: 16,200–2,543,150) to the baseline model total need of 68,139,600 (average per country of 2,839,150 cartridges per year; range: 69,600–25,192,700), with the percentage increase per country ranging from −48% (Tajikistan) to 6706% (Democratic Republic of the Congo).

Under the baseline model, a more than 6-fold increase in the volume of test cartridges would be needed.

Sensitivity and Scenario Analysis

The results of a sensitivity and scenario analysis are presented in Table 4. The full data for this analysis are listed by country in a Supplement. The changes that would, on average, decrease the number of modules needed include the scenarios that use 2017 actual TB case notifications instead of 90% of the estimated burden (leading to a total reduction of 32%), test a lower percentage of PLHIV on ART (leading to a total reduction of 6%–10%), assume a higher module operating capacity (48% reduction), or assume availability and use of CXR to triage persons with symptoms of TB before the WRD (50% reduction). The scenarios resulting in an increase in the number of modules needed above the baseline model include countries with lower TB prevalence, more ambitious TB case finding (with up to 100% or more increase in module needs from baseline), and more ambitious TB screening of PLHIV (6% increase in modules needed). The likelihood of these various scenarios is explored further in the Discussion.

The numbers of WRD sites to mimic the 1 microscope per 100,000 population requirement used for smear microscopy access are a total of 78% (range 30%–92%) smaller than the baseline

TABLE 3. Xpert Modules and Test Cartridges Needed Under Baseline Model for 24 High-Burden Countries, by Country

Country	Current Number of Modules	Total Modules Needed	Change Needed from Current (%)	No. Test Cartridges (2017)	No. Test Cartridges Needed, Annually	Change Needed from Current (%)
Afghanistan	180	1,146	537%	17,500	577,600	3201%
Bangladesh	860	6,229	624%	341,900	3,139,400	818%
Cambodia	300	898	199%	134,050	452,700	238%
Democratic Republic of the Congo	614	4,368	611%	32,350	2,201,700	6706%
Ethiopia	1268	3,111	145%	203,950	1,567,700	669%
India	4780	49,986	946%	2,543,150	25,192,700	891%
Indonesia	2356	14,545	517%	507,450	7,330,800	1345%
Kenya	838	2,914	248%	450,450	1,468,900	226%
Kyrgyzstan	80	167	108%	16,200	84,000	419%
Malawi	428	873	104%	62,150	440,100	608%
Mozambique	368	3,061	732%	150,250	1,542,900	927%
Myanmar	367	3,120	750%	41,300	1,572,400	3707%
Nigeria	1576	8,182	419%	349,850	4,123,500	1079%
Pakistan	2808	9,087	224%	435,050	4,579,900	953%
Philippines	1436	10,029	598%	301,200	5,054,500	1578%
South Africa	4204	7,027	67%	2,198,000	3,541,500	61%
Tajikistan	122	138	13%	133,150	69,600	−48%
Tanzania	852	2,866	236%	395,250	1,444,400	265%
Uganda	994	1,762	77%	300,850	888,000	195%
Ukraine	292	725	148%	80,000	365,200	357%
Uzbekistan	208	453	118%	76,100	228,100	200%
Vietnam	690	2,230	223%	219,500	1,124,000	412%
Zambia	720	1,190	65%	166,850	599,700	259%
Zimbabwe	532	1,092	105%	247,900	550,300	122%
TOTAL	26,873	135,198	403%	9,404,400	68,139,600	625%

values for number of modules, indicating that many of the sites would likely have more than a single module to achieve the baseline scenario (Supplement). Indeed, the number of these sites required under the 1 microscope per 100,000 population calculation is only 13% more, on average, than the current number of modules in these countries (range −87% in South Africa to 180% in India).

Cost Analysis

The cost implications of the baseline model are presented in Table 5. The baseline model would require a total incremental investment across the 24 countries of US$473,920,210 (average per country of $19,746,675; range $70,417–$197,774,132) in Xpert instruments, and $586,177,296 per year in cartridges (average per country of $24,424,054; range −$634,229–$226,042,059), for a total incremental investment of $1,060,097,504 (average per country of $44,170,729; range −$563,812 to $423,816,641). This represents a 5-fold increase over current investment in these countries because the current investment (all machines procured to date, plus the number of cartridges procured in 2017) is

The baseline model would require a 5-fold increase in total investment.

TABLE 4. Sensitivity Analysis Relative to the Baseline Model

Scenario Name	Parameter to Change	Default Value	Effect of Variation on Model Output
1. Current notifications	Estimated TB burden	90%	If estimated TB burden is reduced to show only the capacity needed for current TB notifications, this will reduce the number of modules needed by a total of 32% (range: 8%–55%).
2. Reduced TB prevalence	Number of adults with symptoms needed to test with rapid diagnostic (Xpert) to diagnose 1 adult TB patient (NNT)	10	Number of needed modules changes almost proportionately (e.g., increasing to 12 tests will increase output by up to 19%). As prevalence decreases, the value will increase.
3. More screening of PLHIV	Percentage of PLHIV on ART with signs and symptoms of TB that require Xpert test	20%	Increasing the percentage to 30% will increase number of modules needed by a total of 6% (range: 0%–38%).
4. Less screening of PLHIV			Decreasing to 10% will decrease number of modules needed by a total of 6% (range: 0%–38%).
5. Current screening of PLHIV			Decreasing to 2.5% will decrease number of modules needed by an average of 10% (range: 0%–67%).
6. Increased operation of module	Operational capacity	168 days/year 3 cycles/day	Increasing working days to 240 and test throughput to 4 cycles/day will reduce number of modules needed by 48% in all countries.
7. CXR triage	Number needed to test	10	Including CXR as a triage tool before the WRD is estimated to reduce the number needed to test to 5 for HIV-negative adults and 2 for children, and to reduce the baseline number of PLHIV on ART requiring Xpert testing by 50%; in total, this would therefore reduce the number of modules and cartridges needed by 50%. See text for justification.
8. Ambitious case finding	Number needed to test and operational capacity	10 for NNT; 3 cycles/day	To detect all people with TB, more ambitious case finding is needed. This is likely to result in both more down-time for modules (due to greater decentralization and/or using mobile screening, thus cycles/day is reduced to 2) and a lower positivity rate from testing more people with symptoms of TB (thus NNT is increased to 20). This combination of changes increases modules and cartridges needed by 177% and 84% (range: 80%–194% and 20%–96%).
9. WRD sites	Access standard for smear microscopy	None	Converts access standard for smear microscopy (1 microscope/100,000 population) to WRD sites needed to achieve same geographical coverage. This produces a large number of sites needed, though these values are more than 4 times lower than the baseline modules needed, since each site will require multiple modules to achieve sufficient throughput.

Abbreviations: CXR, chest X-ray; NNT, Number needed to test; PLHIV, people living with HIV; TB, tuberculosis; WRD, World Health Organization-recommended rapid diagnostic.

TABLE 5. Cost Implications of Baseline Model (All Values in US$)

Country	A: Cost of Additional Xpert Modules (Based on the Price of a 4-Module Instrument)	B: Cost of 1 Year's Supply of Additional Cartridges, According to Total Calculated Need	C: Total Incremental Cost of Baseline Model Over Current Situation (New Modules Plus 1 Year of Cartridges) (A+B)	D: Total Current Cost (Existing Modules Plus Number of Cartridges Procured in 2017)	E: Total % Increase in Investment Needed (C/D)
Afghanistan	$4,226,389	$5,589,798	$9,816,187	$962,150	1020%
Bangladesh	$23,489,236	$27,919,050	$51,408,286	$7,174,662	717%
Cambodia	$2,617,188	$3,180,127	$5,797,315	$2,650,319	219%
Democratic Republic of the Congo	$16,425,729	$21,650,113	$38,075,842	$3,009,103	1265%
Ethiopia	$8,061,007	$13,610,225	$21,671,232	$7,582,921	286%
India	$197,774,132	$226,042,509	$423,816,641	$46,293,137	916%
Indonesia	$53,327,917	$68,097,033	$121,424,950	$15,371,851	790%
Kenya	$9,084,618	$10,164,131	$19,248,749	$8,161,741	236%
Kyrgyzstan	$379,167	$676,644	$1,055,811	$511,676	206%
Malawi	$1,947,813	$3,771,941	$5,719,754	$2,492,757	229%
Mozambique	$11,783,229	$13,898,647	$25,681,876	$3,109,495	826%
Myanmar	$12,043,681	$15,280,378	$27,324,059	$2,017,799	1354%
Nigeria	$28,899,271	$37,661,027	$66,560,298	$10,386,503	641%
Pakistan	$27,471,076	$41,365,603	$68,836,679	$16,626,799	414%
Philippines	$37,593,368	$47,437,934	$85,031,302	$9,288,476	915%
South Africa	$12,349,688	$13,408,130	$25,757,818	$40,328,540	64%
Tajikistan	$70,417	$(634,229)	$(563,812)	$1,862,587	-30%
Tanzania	$8,810,694	$10,470,517	$19,281,211	$7,672,095	251%
Uganda	$3,359,583	$5,859,757	$9,219,340	$7,351,233	125%
Ukraine	$1,892,639	$2,846,296	$4,738,935	$2,075,900	228%
Uzbekistan	$1,070,035	$1,516,960	$2,586,995	$1,669,478	155%
Viet Nam	$6,738,194	$9,026,910	$15,765,104	$5,209,360	303%
Zambia	$2,055,729	$4,319,843	$6,375,572	$4,815,163	132%
Zimbabwe	$2,449,410	$3,017,952	$5,467,362	$4,801,542	114%
TOTAL	$473,920,208	$586,177,296	$1,060,097,504	$211,425,287	
AVERAGE	$19,746,675	$24,424,054	$44,170,729	$8,809,387	474%

approximately $211,425,287 (average per country of $8,809,386; range $511,676–$46,293,137).

DISCUSSION

The TB diagnostics network is the foundation for all other interventions needed to end the global TB epidemic. Without an accessible, quality network of rapid TB diagnostics, countries and the global TB community will never reach the 40 by 2022 goal.

We did an estimation exercise to determine if countries have adequate rapid TB testing capacity to be able to detect all people with TB, in line with the 40 by 2022 goal. This analysis showed that there is a considerable gap between the existing rapid testing capacity and the capacity that is actually needed. Compared to the current situation, the baseline model required a 4-fold expansion in Xpert module capacity across these 24 countries with high TB burdens, and a 6-fold increase in

Without an accessible, quality network of rapid TB diagnostics, countries and the global TB community will never reach the 40 by 2022 goal.

the number of test cartridges. The estimated total cost for this scale-up is approximately $474 million for additional modules, plus an incremental cost each year of cartridges of approximately $586 million. Proportionately, even greater needs are possible in countries with more active case finding, more aggressive screening of PLHIV, usage of Xpert machines for other diseases,[37] or lower TB prevalence (because countries would need more tests to find the same number of TB cases).

These large volumes in the baseline model are consistent with early predictions of substantial potential market sizes for new TB diagnostics, with the prediction that 59% adoption of a smear-replacement test by 2020 would result in an estimated annual volume of 49 million tests.[38] By comparison, the actual procurement reported in 2017 by Cepheid across all high-TB burden countries was less than a quarter of that, at 11.2 million tests procured (Philippe Jacon, Cepheid, email communication, November 2018).

For various reasons, a total required volume lower than that identified from the baseline model is possible (Table 4), but not likely. First, the ratio of people with TB symptoms to diagnosed TB patients may be less than 10:1, though some of the best data for this come from South Africa and point to, if anything, a higher number. Second, using the national TB program as the data source for Xpert module numbers may miss modules that have been procured directly by the private sector with their own private funding. However, the numbers of such machines in TB high-burden countries are minimal (and zero in many high-burden countries).[39] Third, the use of CXR as a triage tool before the WRD could potentially approximately halve the WRD needs, though cost and major patient access issues around CXR have resulted in limited usage of such an algorithm.

Indeed, it may be optimistic to estimate an approximate 50% reduction in Xpert testing volume based on adding a CXR triage step. In the TB prevalence survey in Vietnam, 3.7% of the general population had a CXR abnormality,[40] in a population where TB prevalence was 260 per 100,000 (0.26%), so the ratio of CXR abnormality to confirmed TB was greater than 10 to 1 and thus the NNT in a CXR triage algorithm remained high. Practicality of this algorithm is also a concern. In terms of resource needs, as a first approximation we could assume the same number of CXR machines being needed as the number of Xpert sites under the "per 100,000 population" calculation (see Supplement). Because a CXR takes only a few minutes, throughput reasons would likely not justify a greater number of CXR machines than Xpert machines. However, for patient accessibility, such an assumption is very much on the low end because it is far easier to transport sputum to an Xpert than patients to a CXR. Thus, this estimated need for CXR under this scenario should be considered a low-end estimate.

Fourth and finally, operational capacity of WRD machines might improve (less down-time and more cycles per day), though this is not what we see from current experiences in high-burden countries where there is increasing evidence for "Xpert for all" algorithms being incompletely implemented due to resource constraints. For example, Ethiopia has been aggressive in adopting an "Xpert for all" algorithm, which has almost tripled the use of Xpert in 3 years, but the peak utilization is still only 93%[41] of the "realistic operating capacity" defined above (see Methods) or 54% of the original WHO implementation recommendations (3–4 tests/module/day · 250 days per year).[42] This less than optimal operating capacity is not because of a lack of need; the percentage of TB cases tested for rifampicin resistance in these project areas has increased but is still only 28%.[41] Indonesia represents another example, where substantial support for an "Xpert for all" algorithm in focus districts has resulted in more than double the use of Xpert, but peaking at only 38% of full operational capacity in focus districts (compared to 16% in non-supported provinces, using the WHO implementation recommendations).[43] Thus, even with substantial support and an expansive algorithm, the instrument's maximum potential operating capacity seems out of reach and the capacity presented in the baseline model presents a more realistic scenario.

Two other ways to reach lower numbers for the resource needs would be to assume either current case finding numbers (instead of the targeted 90% of total incident cases) and the current, inadequate TB screening percentages for PLHIV. Although in Table 4 we present the results of using such sensitivity analyses, our baseline model was aimed explicitly at estimating the future needs of the TB community to reach goals that have already been set. Both of these 2 sensitivity analyses go against that theme by settling for the status quo. Thus, although these 2 analyses are included in Table 4 for the sake of completeness, we do not see them as a challenge to the baseline numbers.

It is important to note that the number of Xpert tests needed for PLHIV screening does not decrease greatly with the widespread adoption of ART. Even PLHIV on ART will still have significant

levels of TB symptoms from non-TB causes, just like the general population, and it is these non-TB causes that are behind the vast majority of the symptoms that prompt an Xpert test. The subsequent TB yield from those Xpert tests will decrease for people on ART, but the need for the tests in the first place remains, which is the relevant issue for this exercise.

When Xpert was introduced in 2010, it was intended to be a point-of-care test to replace smear microscopy. By doing this, countries would keep access intact but significantly improve sensitivity and be able to diagnose rifampicin resistance with the initial test. In reality, roll-out has been steady but slow.[44,45] In 2013, WHO policy[46] recommended that Xpert be used as the initial diagnostic test in adults and children suspected of having multidrug-resistant TB or HIV-associated TB; the use of Xpert as the initial test in all adults and children with symptoms of TB was a conditional recommendation and not taken up by most countries. Although the WHO Compendium[8] states that Xpert is to be used as the initial diagnostic test in everyone with symptoms of TB (see Methods), the original restricted policy led to an uneven and slow uptake of Xpert as the primary diagnostic for all people with symptoms of TB. By the end of 2017, national algorithms and policies in only 32 of the 48 countries included in WHO's lists of high TB, TB/HIV, and multidrug-resistant TB burden countries had been revised to include this recommendation for use of Xpert for all individuals with TB symptoms,[3] and the extent of implementation in countries with these policies varies.

Why are WRDs including Xpert not being used universally as the primary diagnostic tool for TB? Beyond the explanation of insufficient financial resources, there are several plausible reasons.[47] Originally, Xpert was intended to be a "near" point-of-care diagnostic placed within subdistrict facilities similar to the level of smear microscopy services; however, limited resources and operational challenges like unstable power supply forced Xpert to become more centralized and ultimately inaccessible without specimen transport mechanisms in place.[48] Some of the operational barriers[48,49] are slowly being addressed, including the use of alternative power sources like solar and diagnostic data management solutions like GXAlert/ASPECT that provides visibility to the program on all instruments (thus enabling a response to instrument problems or commodity issues). There is also an abundance of training material for all levels of the health system that can be used to build staffing capacity and there have been creative approaches to address staff shortages. But, despite all these interventions, the ability to move Xpert to the level of the microscopy center is still in doubt in many high-burden countries. Issues such as power, infrastructure (i.e., air conditioning), capacity of staff to troubleshoot, lack of maintenance and service, and module failures remain major operational challenges that will have to be faced in any ongoing expansion, with a focus on supporting the systems required for a true point-of-care functionality.[50,51] In addition, since concessional pricing for the Xpert instrument and test cartridges is limited to the public sector, the test is mostly unavailable to persons who seek care from private providers and facilities.[39]

Limitations

Our analysis has a number of limitations. There is limited information to inform the setting of values for certain key variables, including NNT. True numbers will, in any case, vary substantially between countries depending on epidemiology, the intensity of case finding, and other factors. Attempts to incorporate such considerations via stratification of the model would result in a less transparent and more questionable model based on suppositions rather than evidence and was therefore not undertaken. However, we believe that the baseline model remains a reasonable estimate that errs on the side of conservatism. For the estimation of current cartridge procurement volumes, the estimated values may not reflect true consumption if the procurement order for a country is not based on the previous year's consumption due to leftover stock from the previous year or an increase or decrease in funding available for procurement. In addition, the resource needs estimate does not include a number of additional and substantial areas of investment that would be needed including the cost of maintenance contracts, shipping and customs payments, connectivity installment and maintenance, sputum transportation, infrastructure requirements, and training and paying salaries for additional staff. Finally, the required investment amount may differ if different WRDs, with a different price, are used instead of Xpert,[14] or if Xpert machine or cartridge procurement is by private providers without access to concessional pricing.

CONCLUSION

Even as countries continue to work out WRD expansion and operational issues, the issue of the

total capacity needed (as addressed here) also remains, including the gap in the resources needed to reach that capacity. Rapid test availability is very far from the only issue and need that is confronting TB programs as they aim for the 40 by 2022 targets; there are a multitude of additional activity, financial, and system constraints that must also be addressed. However, it is clear from this analysis that countries do not have enough rapid TB test instruments or cartridges to meet their needs. Without increasing both instrument and cartridge numbers, countries will struggle to find all people with TB and to implement quality TB diagnostics at scale. Ambitious goals such as the 40 by 2022 require bold interventions. This includes urgently expanding access to and capacity of country TB diagnostic testing networks.

Acknowledgments: The authors would like to thank C. Vincent, S. Ahmedov, P. Pierre, and the USAID/Washington TB team for data sources and comments on the analysis. Additionally, we thank our colleagues in USAID missions for the 24 countries presented in the analysis for sourcing the detailed data on current numbers of Xpert modules.

The authors are employed by the United States Agency for International Development (USAID), Washington DC, USA. The views and opinions expressed in this article are those of the authors and not necessarily the views and opinions of USAID.

Funding: No specific funding for this work.

Competing Interests: None declared.

REFERENCES

1. United Nations High-Level Meeting on the Fight to End Tuberculosis web page. https://www.un.org/pga/73/event-detail/fight-to-end-tuberculosis/. Accessed July 1, 2019.
2. Resolution adopted by United Nations General Assembly on 10 October 2018. Political declaration of the High-Level Meeting of the General Assembly on the Fight Against Tuberculosis. https://digitallibrary.un.org/record/1649568/files/A_RES_73_3-EN.pdf. Accessed July 1, 2019.
3. World Health Organization (WHO). *Global Tuberculosis Report, 2018*. Geneva, Switzerland: WHO; 2018. http://apps.who.int/iris/bitstream/handle/10665/274453/9789241565646-eng.pdf. Accessed July 1, 2019.
4. Walzl G, McNerney R, du Plessis N, et al. Tuberculosis: advances and challenges in development of new diagnostics and biomarkers. *Lancet Infect Dis*. 2018;18(7):e199–e210. CrossRef. Medline
5. Steingart KR, Ramsay A, Pai M. Optimizing sputum smear microscopy for the diagnosis of pulmonary tuberculosis. *Expert Rev Anti Infect Ther*. 2007;5(3):327–331. CrossRef. Medline
6. Perkins MD, Roscigno G, Zumla A. Progress towards improved tuberculosis diagnostics for developing countries. *Lancet*. 2006;367(9514):P942–P943. CrossRef. Medline
7. World Health Organization (WHO). *The End TB Strategy*. Geneva, Switzerland: WHO; 2015. http://www.who.int/tb/End_TB_brochure.pdf. Accessed July 1, 2019.
8. World Health Organization (WHO). *Compendium of WHO Guidelines and Associated Standards: Ensuring Optimum Delivery of the Cascade of Care for Patients with Tuberculosis*. 2nd ed. Geneva, Switzerland: WHO; 2018. http://apps.who.int/iris/bitstream/handle/10665/272644/9789241514101-eng.pdf. Accessed July 1, 2019.
9. Global Laboratory Initiative. *GLI Model TB Diagnostic Algorithms*. Geneva, Switzerland: WHO; 2018. http://www.stoptb.org/wg/gli/assets/documents/GLI_algorithms.pdf. Accessed July 1, 2019.
10. World Health Organization. Global Tuberculosis Database. Geneva, Switzerland: WHO; 2019. http://www.who.int/tb/country/data/download/en/. Accessed July 1, 2019.
11. World Health Organization. *Systematic Screening for Active Tuberculosis: Principles and Recommendations*. Geneva, Switzerland: WHO; 2013. https://www.who.int/tb/publications/Final_TB_Screening_guidelines.pdf. Accessed July 1, 2019.
12. World Health Organization. *Systematic Screening for Active Tuberculosis: An Operational Guide*. Geneva, Switzerland: WHO; 2015. http://apps.who.int/iris/bitstream/handle/10665/181164/9789241549172_eng.pdf. Accessed July 1, 2019.
13. Wells WA. Onions and prevalence surveys: how to analyze and quantify tuberculosis case-finding gaps. *Int J Tuberc Lung Dis*. 2017;21(11):1101–1113. CrossRef. Medline
14. Unitaid. *Tuberculosis: Diagnostics Technology and Market Landscape*. 5th ed. Geneva, Switzerland: Unitaid; 2017.
15. World Health Organization. *Spreadsheet for Country-Specific Calculation of Diagnostic Testing and Facility Needs Using Recommended Algorithms*. http://www.who.int/tb/publications/calculations_of_lab_capacity.xls. Accessed July 1, 2019.
16. United States Agency for International Development (USAID). USAID TB Priority Countries. https://www.usaid.gov/global-health/health-areas/tuberculosis/countries. Accessed July 1, 2019.
17. Joint United Nations Programme on HIV/AIDS. *UNAIDS Data 2017*. Geneva, Switzerland: UNAIDS; 2017. http://www.unaids.org/sites/default/files/media_asset/20170720_Data_book_2017_en.pdf. Accessed July 1, 2019.
18. Joint United Nations Programme on HIV/AIDS (UNAIDS). *90-90-90: An Ambitious Treatment Target to Help End the AIDS Epidemic*. Geneva, Switzerland: UNAIDS; 2017. http://www.unaids.org/en/resources/909090. Accessed July 1, 2019.
19. Stop TB Partnership and World Health Organization. *The Global Plan to Stop TB 2011-2015*. Geneva, Switzerland: WHO; 2010. http://www.stoptb.org/assets/documents/global/plan/tb_globalplantostoptb2011-2015.pdf. Accessed July 1, 2019.
20. Accessible pricing. Foundation for Innovative New Diagnostics website. https://www.finddx.org/pricing/. Accessed July 1, 2019.
21. World Health Organization (WHO). *Laboratory Services in Tuberculosis Control: Organization and Management, Part 1*. Geneva, Switzerland: WHO; 1998. http://apps.who.int/iris/bitstream/handle/10665/65942/WHO_TB_98.258_(part1).pdf. Accessed July 1, 2019.
22. World Health Organization (WHO). *Framework of Indicators and Targets for Laboratory Strengthening Under the End TB Strategy*. Geneva, Switzerland: WHO; 2016. http://apps.who.int/iris/bitstream/10665/250307/1/9789241511438-eng.pdf. Accessed July 1, 2019.
23. Stevens WS, Scott L, Noble L, Gous N, Dheda K. Impact of the GeneXpert MTB/RIF technology on tuberculosis control. *Microbiol Spectr*. 2017;5(1). CrossRef. Medline
24. National Health Laboratory Service (NHLS). *GeneXpert MTB/RIF Progress Report, July 2014*. Johannesburg, South Africa: NHLS; 2014. http://www.nhls.ac.za/wp-content/uploads/2018/09/GeneXpert-Progress-Report-July-2014_Final.pdf. Accessed July 1, 2019.
25. National Health Laboratory Service (NHLS). *GeneXpert MTB/RIF National Report, July 2018*. Johannesburg, South Africa: NHLS; 2018.

26. PEPFAR. PEPFAR data dashboards. President's Emergency Plan for AIDS Relief website. https://data.pepfar.gov/. Accessed July 1, 2019.
27. World Health Organization. *Guidelines for Intensified Tuberculosis Case-Finding and Isoniazid Preventive Therapy for People Living with HIV in Resource-Constrained Settings*. Geneva, Switzerland: WHO; 2011. http://apps.who.int/iris/bitstream/handle/10665/44472/9789241500708_eng.pdf. Accessed July 1, 2019.
28. Getahun H, Kittikraisak W, Heilig CM, et al. Development of a standardized screening rule for tuberculosis in people living with HIV in resource-constrained settings: individual participant data meta-analysis of observational studies. *PLoS Med*. 2011;8(1):e1000391. CrossRef. Medline
29. Ohene SA, Bonsu F, Hanson-Nortey NN, et al. Provider initiated tuberculosis case finding in outpatient departments of health care facilities in Ghana: yield by screening strategy and target group. *BMC Infect Dis*. 2017;17(1):739. CrossRef. Medline
30. Padmapriyadarsini C, Bhavani PK, Sekar L, et al. Effectiveness of symptom screening and incidence of tuberculosis among adults and children living with HIV infection in India. *Natl Med J India*. 2016; 29(6):321–325. Medline
31. Bjerrum S, Bonsu F, Hanson-Nortey NN, et al. Tuberculosis screening in patients with HIV: use of audit and feedback to improve quality of care in Ghana. *Glob Health Action*. 2016;9(1):32390. CrossRef. Medline
32. Turinawe K, Vandebriel G, Lowrance DW, et al. Operating characteristics of a tuberculosis screening tool for people living with HIV in out-patient HIV care and treatment services, Rwanda. *PLoS One*. 2016;11(9):e0163462. CrossRef. Medline
33. Adelman MW, Tsegaye M, Kempker RR, et al. Intensified tuberculosis case finding among HIV-infected persons using a WHO symptom screen and Xpert® MTB/RIF. *Int J Tuberc Lung Dis*. 2015;19(10): 1197–1203. CrossRef. Medline
34. Modi S, Cavanaugh JS, Shiraishi RW, et al. Performance of Clinical Screening Algorithms for Tuberculosis Intensified Case Finding among People Living with HIV in Western Kenya. *PLoS One*. 2016;11(12):e0167685. CrossRef. Medline
35. World Health Organization (WHO). *Chest Radiography in Tuberculosis Detection: Summary of Current WHO Recommendations and Guidance on Programmatic Approaches*. Geneva, Switzerland: WHO; 2016. http://apps.who.int/iris/bitstream/handle/10665/252424/9789241511506-eng.pdf. Accessed July 1, 2019.
36. van't Hoog AH, Meme HK, Laserson KF, et al. Screening strategies for tuberculosis prevalence surveys: the value of chest radiography and symptoms. *PLoS One*. 2012;7(7):e38691. CrossRef. Medline
37. Cazabon D, Pande T, Kik S, et al. Market penetration of Xpert MTB/RIF in high tuberculosis burden countries: A trend analysis from 2014—2016. *Gates Open Res*. 2018;2:35. CrossRef. Medline
38. World Health Organization/Special Programme for Research and Training in Tropical Diseases/Foundation for Innovative New Diagnostics. *Diagnostics for Tuberculosis: Global Demand and Market Potential*. Geneva, Switzerland: WHO; 2006. http://apps.who.int/iris/handle/10665/43543. Accessed July 1, 2019.
39. Puri L, Oghor C, Denkinger CM, Pai M. Xpert MTB/RIF for tuberculosis testing: access and price in highly privatised health markets. *Lancet Glob Health*. 2016;4(2):e94–e95. CrossRef. Medline
40. Hoa NB, Cobelens FGJ, Sy DN, Nhung NV, Borgdorff MW, Tiemersma EW. Yield of interview screening and chest X-ray abnormalities in a tuberculosis prevalence survey. *Int J Tuberc Lung Dis*. 2012;16(6):762–767. CrossRef. Medline
41. Challenge TB. *Challenge TB-Ethiopia: Year 4 Annual Report*. Addis Ababa, Ethiopia: Challenge TB; 2018. https://pdf.usaid.gov/pdf_docs/PA00TDJM.pdf. Accessed July 1, 2019.
42. World Health Organization. *Xpert MTB/RIF Implementation Manual: Technical and Operational 'How-to'; Practical Considerations*. Geneva, Switzerland: WHO; 2014. http://apps.who.int/iris/bitstream/handle/10665/112469/9789241506700_eng.pdf. Accessed July 1, 2019.
43. Challenge TB. Challenge TB-Indonesia: Year 4 Annual Report. Jakarta: Indonesia: Challenge TB; 2018. https://pdf.usaid.gov/pdf_docs/PA00TDJN.pdf. Accessed July 1, 2019.
44. Qin ZZ, Pai M, Van Gemert W, Sahu S, Ghiasi M, Creswell J. How is Xpert MTB/RIF being implemented in 22 high tuberculosis burden countries? *Eur Respir J*. 2015;45(2):549–554. CrossRef. Medline
45. Cazabon D, Suresh A, Oghor C, et al. Implementation of Xpert MTB/RIF in 22 high tuberculosis burden countries: are we making progress? *Eur Respir J*. 2017;50(2):1700918. CrossRef. Medline
46. World Health Organization (WHO). Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF assay for the diagnosis of pulmonary and extrapulmonary TB in adults and children. Policy update. Geneva, Switzerland: WHO; 2013. http://apps.who.int/iris/bitstream/handle/10665/112472/9789241506335_eng.pdf. Accessed July 1, 2019.
47. Lawn SD, Mwaba P, Bates M, et al. Advances in tuberculosis diagnostics: the Xpert MTB/RIF assay and future prospects for a point-of-care test. *Lancet Infect Dis*. 2013;13(4):349–361. CrossRef. Medline
48. Albert H, Nathavitharana RR, Isaacs C, Pai M, Denkinger CM, Boehme CC. Development, roll-out and impact of Xpert MTB/RIF for tuberculosis: what lessons have we learnt and how can we do better? *Eur Respir J*. 2016;48(2):516–525. CrossRef. Medline
49. Pai NP, Vadnais C, Denkinger C, Engel N, Pai M. Point-of-care testing for infectious diseases: diversity, complexity, and barriers in low- and middle-income countries. *PLoS Med*. 2012;9(9):e1001306. CrossRef. Medline
50. Pai M, Pai NP, Ghiasi M. Point-of-Care Diagnostic Testing in Global Health: What Is the Point? *Microbe Magazine*. 2015;10(3):103–107. CrossRef
51. García-Basteiro AL, DiNardo A, Saavedra B, et al. Point of care diagnostics for tuberculosis. *Pulmonology*. 2018;24(2):73–85. CrossRef. Medline

Peer Reviewed

Received: July 2, 2019; **Accepted:** September 17, 2019; **First Published Online:** December 5, 2019

Cite this article as: Piatek AS, Wells WA, Shen KC, Colvin CE. Realizing the "40 by 2022" commitment from the United Nations High-Level Meeting on the Fight to End Tuberculosis: what will it take to meet rapid diagnostic testing needs? *Glob Health Sci Pract*. 2019;7(4):551-563. https://doi.org/10.9745/GHSP-D-19-00244

GLOBAL HEALTH: SCIENCE AND PRACTICE
Dedicated to what works in global health programs
OPEN ACCESS

ORIGINAL ARTICLE

Standardizing Measurement of Contraceptive Use Among Unmarried Women

Madeleine Short Fabic,[a] Apoorva Jadhav[a]

Because contraceptive prevalence and unmet need for family planning estimates for unmarried women vary widely depending on the chosen sexual recency inclusion factor, all data platforms should adopt a 1-month window in these calculations to have comparable and actionable estimates.

ABSTRACT

Historically, the family planning practices and needs of married women have been monitored and reported uniformly. However, the same uniformity does not hold for unmarried women. Because key data and information platforms employ different measurement approaches—namely, different definitions of sexual recency—reports of contraceptive prevalence and unmet need among unmarried women are inconsistent. We examine how the measurement approaches employed by 3 large organizations yield such divergent estimates. We find that contraceptive prevalence and unmet need estimates among married women do not vary much by sexual recency. For unmarried women, contraceptive prevalence is systematically lower and unmet need is systematically higher as the sexual recency window widens. In the short term, we recommend using the 1-month cutoff as analyses reveal it yields the most precise estimates for better recognizing the needs of this important demographic group.

BACKGROUND

With the world working to achieve the FP2020 goal that 120 million more women and girls access contraception and the 2030 Sustainable Development Goal that no one is left behind, understanding the family planning needs of all women is imperative.[1] The contraceptive prevalence of married and in-union women has been monitored and reported in a relatively uniform way across time, place, and data collection instrument.[2,3] The same cannot be said for the monitoring and reporting of family planning practices and needs among unmarried women.

When World Fertility Surveys began in the 1960s and 1970s, the approach to measuring contraceptive prevalence only among married and in-union women had merit due to cultural sensitivities of asking unmarried women about their sexual activity coupled with assumptions of low sexual activity among unmarried women.[4] During the 1980s, this approach shifted, and it became standard to include unmarried women in population-based household surveys and other data collection platforms. This expansion was and remains important.

Although marriage is still largely the norm worldwide—according to the United Nations World Marriage Data, at least 80% of women aged 45–49 have ever married[5]—marriage is occurring later in life. Meanwhile, age at sexual debut remains relatively constant.[5,6] Coupled with large youth populations in many low- and middle-income countries, this demographic shift yields not only a higher proportion of unmarried women and men exposed to the risk of pregnancy but also the highest number of unmarried individuals in human history.[7] The increased focus on expanding family planning access among youth,[8] who include a significant share of unmarried individuals, is clearly warranted. And the inclusion of unmarried women and men in key data collection platforms is imperative for understanding the family planning knowledge, attitudes, and practices among this growing demographic.

The key family planning data collection platforms—Demographic and Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS), and Performance Monitoring and Accountability 2020 surveys—have built on lessons learned over the decades, including the experiences of the World Fertility Surveys[9] and Contraceptive Prevalence Surveys.[10] All 3 platforms include both married and unmarried women, except in countries (largely in North Africa and South and West Asia) where only married women are included due to cultural sensitivities, and ask questions on contraceptive use and sexual activity regardless of marital status.[11]

The calculation and reporting of key family planning indicators for married women is uniform across

[a] United States Agency for International Development, Washington, DC, USA. Correspondence to Madeleine Short Fabic (mshort@usaid.gov).

information platforms, but varies for unmarried women. For example, take 3 key family planning data and communication groups: DHS, the Guttmacher Institute, and World Health Organization (WHO). All 3 groups frequently produce family planning-related reports, data from which are used widely for policy, programmatic, research, and advocacy purposes. DHS reports the modern contraceptive prevalence rate (mCPR) among unmarried women only among those who reported having sex in the 1 month preceding the interview.[12] The Guttmacher Institute and recent WHO publications report mCPR among women who reported having sex in the 3 months preceding the interview.[13,14] This Guttmacher/WHO approach aligns with the research of Dasgupta et al. that found that limiting analysis to 1 month before the interview misses a significant proportion of unmarried women who were sexually active within a wider window of time.[15] Ultimately, these varying approaches to measuring mCPR among unmarried women result in different mCPR estimates as well as different unmet need estimates. These differing estimates have potentially significant policy and programming ramifications (e.g., size of population in need). Given these differences, coupled with the increased need to address family planning among unmarried individuals, we aimed to address the following 3 questions:

1. How do various measurement approaches, namely different durations of sexual recency, impact estimates of mCPR and unmet need among unmarried women?
2. What are the benefits and limitations of each measurement approach?
3. Is it beneficial to standardize the measurement approach? Why or why not?

DATA AND METHODS

Data came from the DHS, nationally representative household surveys that 90 countries have implemented with technical assistance from the DHS Program since 1984, when the U.S. Agency for International Development began the program. All women aged 15–49 in sampled households are eligible for the women's interview, further details of which are elaborated upon elsewhere.[16] Our study included DHS data that met the following 4 inclusion criteria: (1) data were collected from both married and unmarried women; (2) data were collected from 2012 or later (we chose the 2012 demarcation to complement FP2020 goals outlined in the 2012 London Summit); (3) final datasets were available for download and analysis as of February 2019; and (4) for countries that had more than 1 DHS since 2012, we included only the most recent.

Our final study dataset included 43 country surveys across 5 regions: Asia and Pacific, West Asia and Eastern Europe, East and Southern Africa, West and Central Africa, and Latin America and the Caribbean. Survey sample sizes ranged from 5,329 women in Comoros to 699,686 in India.

The key demographic variable in our analysis was marital status, unmarried (never married or formerly married) and currently married (married or in-union). We first ran our analysis by separating women into 3 groups, "never married," "currently married or in-union," and "formerly married (separated, divorced, or widowed)." Due to small sample sizes in the latter group, particularly when comparing sexual recency variables (described below), we decided to combine "formerly married" and "never married" into 1 group, "unmarried," with a majority of women being represented by the "never married" group. Recent research has found that the prevalence of divorce is low, ranging from 1%–20% in sub-Saharan Africa,[17] and the proportions are either stable or declining.[18] In our study countries, the "unmarried" group was dominated by "never married women" who comprised 80% of all "unmarried" women, ranging from 71% in Latin America and Caribbean to 86% in Asia. Although we were aware of meaningful differences in being "never married" and "formerly married," we combined the 2 categories due to power limitations and interpreted our results with that caveat. Additionally, we initially included age as a key demographic variable in our analysis but chose not to present the results herein as they were not particularly meaningful due to very small sample sizes by age group (tables available on request). Finally, we considered whether our marital status grouping aligned with our sexual recency variables (described below); that is, did current marital status align with sexual activity status of 1–12 months before the interview. To ensure we were using comparable categories, we analyzed marital duration at the time of survey (note: DHS does not capture data on timing of divorce). On average, only 4% of women were married for less than 12 months at the time of survey (ranging from 2% to 7% across study countries), and the overwhelming majority of these newlywed women (81%) reported sex in the last month. We decided that the potential misalignment between

Differing estimates of mCPR and unmet need among unmarried women have potentially significant policy and programming ramifications.

sexual activity recency and marital duration recency was too small to skew results; therefore, we consistently categorized "currently married" women as "married or in-union" for each of our sexual recency variables.

Our main analytic variable was timing of last sex (e.g., sexual recency). In DHS, participants are asked how old they were when they first had sexual intercourse. For those who say they have ever been sexually active, they are then asked an open-ended question, "When was the last time you had sexual intercourse?" For women who respond that they have been sexually active within the 12 months preceding the interview, the answer is recorded in days, weeks, or months. We used the variable that asked respondents what their time since last sex was in months, what we call "sexual recency," to form 4 corresponding analytic groups:

1. DHS method (sexually active within the 4 weeks/1 month preceding the interview; note that in this article, we use the term 1 month)
2. Guttmacher Institute/WHO method (sexually active within the 3 months preceding the interview; note that the Guttmacher calculation includes months 0, 1, 2, and 3)
3. An alternative method periodically used in research (sexually active in the 12 months preceding interview)
4. All sexually active women regardless of timing of last sex (e.g., those women who have ever had sex)

We excluded incomplete or inconsistent responses. Additionally, we excluded women for whom responses were flagged for various reporting issues.[19] For example, we treated responses that were "before last birth" as instances where the timing of last sex was more than 12 months prior.

The 2 key family planning variables are the mCPR and unmet need. We aligned our definition of modern contraceptive methods with the DHS definition, which includes female and male sterilization, contraceptive pill, intrauterine device, injectables, implants, male and female condoms, diaphragm, contra-ceptive foam and jelly, lactational amenorrhea method, Standard Days method, and emergency contraception.[12] We also used the DHS definition of mCPR, which assesses contraceptive use based on the question, "Are you or your partner currently doing something or using any method to delay or avoid getting pregnant?" As Dasgupta et al. (2017) describe, the interpretation of "current" is up to the woman, ranging from whether she means today, within the last month, or at the time of last sexual intercourse.[15]

We used the DHS Program's revised calculation of unmet need as a starting point for unmet need estimation.[20] Unmet need is defined as the percentage of women who are not currently using a method of contraception and want to stop or delay childbearing.[20] Typically, calculations of unmet need assume all married women are sexually active and, thus, at risk of unmet need if they are not using contraception and do not desire to become pregnant,[21] but because unmarried women who report no sexual intercourse in the 1 month before the interview are assumed to be unexposed to the risk of pregnancy, they are considered not at risk of having unmet need.[20] Some of these assumptions of risk have been challenged over time, and sensitivity tests demonstrate that unmet need estimates are indeed affected by the assumption used.[22] Because the DHS calculation of unmet need among unmarried women applies the inclusion criteria of sex in the 1 month before the interview, we recalculated unmet need by relaxing the sexual activity assumptions to align with each of our sexual recency categories.

First, we explored how various sexual recency "cutoffs" impacted the proportion of women included in mCPR and unmet need estimates by marital status. These results allowed us to visualize the change in denominator (universe of eligible women) based on timing of sex. Next, we explored how estimates of mCPR and unmet need vary by timing of last sex. We also disaggregated these estimates of mCPR and unmet need by marital status. All data presented were weighted. For regional averages presented, all averages were equally weighted by country. We used Stata 14 for data analysis.

RESULTS

How Do Various Sexual Recency Cutoffs Impact the Proportion of Women Included in mCPR and Unmet Need Estimates?

The Table presents data on the variation in the number and percentage of eligible women based on the timing of sexual recency, by region and marital status. Overall, the majority of women report ever having sexual intercourse, ranging from an average of 70% of all women in the Asia/Pacific region to 83% in the West and Central Africa region. These percentages markedly decrease across

TABLE. Mean Percentage and Mean Number of Women Reporting Sex Within Various Time Frames, by Region and Marital Status (as Averaged by Country)

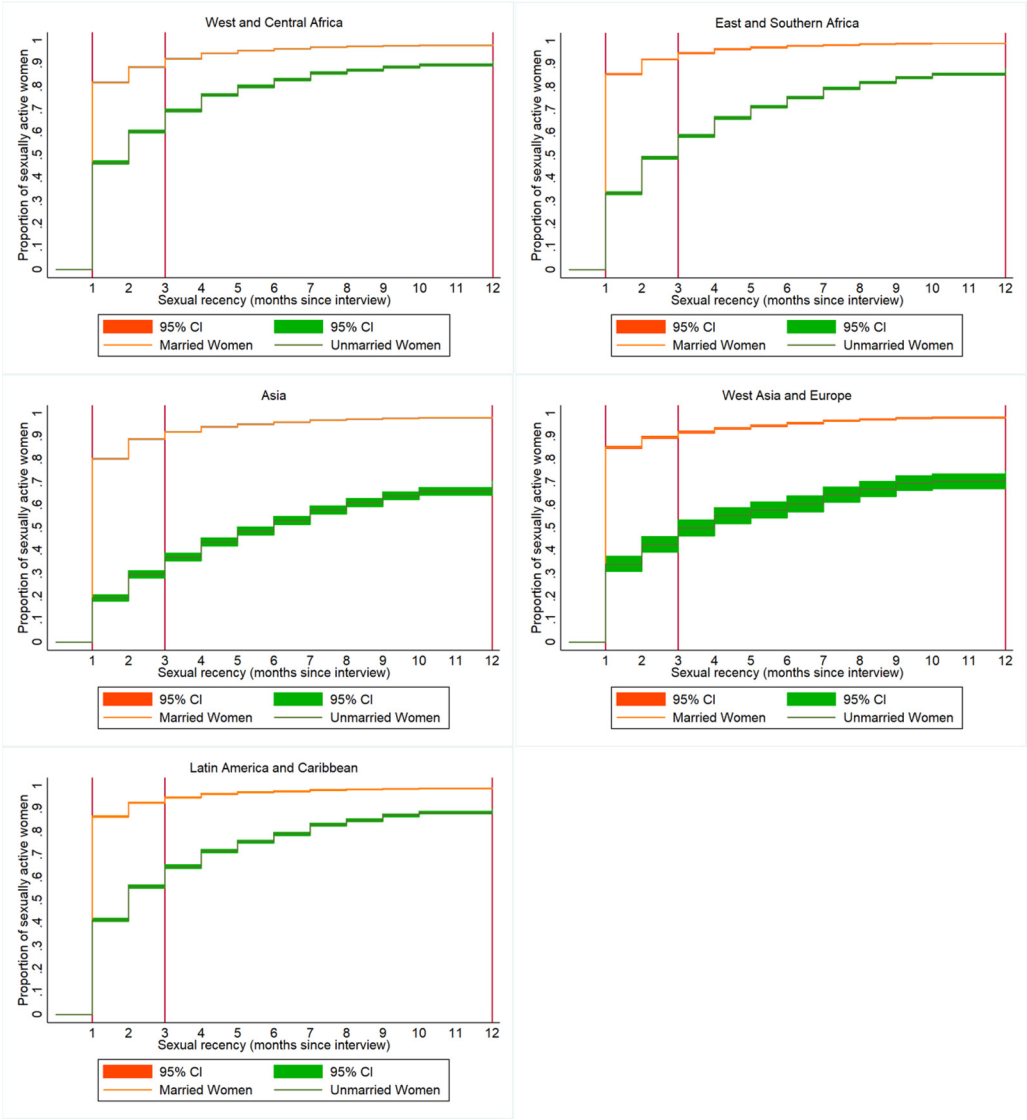
	All Women				
	Total % (No.)	Ever Had Sex % (No.)	Sex in Last 12 Months % (No.)	Sex in Last 3 Months % (No.)	Sex in Last 1 Month % (No.)
East and Southern Africa	100.0 (14,589)	79.2 (11,650)	64.5 (9,000)	57.4 (8,055)	46.4 (6,603)
West and Central Africa	100.0 (14,022)	82.3 (11,451)	69.4 (9,724)	62.7 (8,866)	51.1 (7,321)
West Asia/Europe	100.0 (8,976)	73.9 (6,666)	64.4 (5,777)	59.8 (5,333)	54.5 (4,823)
Asia and Pacific	100.0 (118,043)	70.2 (83,349)	59.0 (21,541)	54.4 (20,068)	46.3 (17,006)
Latin America and Caribbean	100.0 (21,409)	82.4 (17,764)	72.8 (15,631)	64.6 (13,987)	53.1 (11,551)
	Unmarried Women				
	Total % (No.)	Ever Had Sex % (No.)	Sex in Last 12 Months % (No.)	Sex in Last 3 Months % (No.)	Sex in Last 1 Month % (No.)
East and Southern Africa	100.0 (6,075)	58.0 (3,591)	36.3 (2,114)	24.2 (1,406)	12.4 (710)
West and Central Africa	100.0 (4,145)	58.2 (2,463)	42.2 (1,807)	32.0 (1,379)	19.5 (847)
West Asia/Europe	100.0 (2,822)	21.4 (609)	6.8 (201)	3.7 (111)	2.2 (69)
Asia and Pacific	100.0 (32,303)	15.9 (3,955)	4.5 (479)	1.9 (210)	0.8 (93)
Latin America and Caribbean	100.0 (9,505)	61.8 (5,964)	46.0 (4,430)	32.5 (3,174)	18.5 (1,838)
	Married Women				
	Total % (No.)	Ever Had Sex % (No.)	Sex in Last 12 Months % (No.)	Sex in Last 3 Months % (No.)	Sex in Last 1 Month % (No.)
East and Southern Africa	100.0 (8,063)	99.9 (8,059)	90.2 (6,885)	86.9 (6,649)	76.3 (5,893)
West and Central Africa	100.0 (8,998)	99.9 (8,988)	86.9 (7,918)	81.6 (7,487)	69.6 (6,474)
West Asia/Europe	100.0 (6,060)	100.0 (6,057)	93.1 (5,576)	87.7 (5,222)	80.5 (4,754)
Asia and Pacific	100.0 (79,507)	100.0 (79,394)	87.4 (21,063)	82.0 (19,858)	70.2 (16,914)
Latin America and Caribbean	100.0 (11,800)	100.0 (11,800)	95.2 (11,201)	91.5 (10,813)	81.8 (9,713)

all 5 regions as the sexual recency time frame narrows to sex within the last 12 months, 3 months, and finally 1 month (Figure 1). Among unmarried women, the percentage of women who report ever having sex varies dramatically by regional context, from 16% in Asia and Pacific to 62% in Latin America and Caribbean. These percentages decline as the sexual recency window narrows. At the 1-month mark, on average only 1% of unmarried women in Asia and Pacific report sexual activity. The highest percentage of unmarried women reporting sex in the last month (19%) is in West and Central Africa.

Currently married women are more likely to have had sex recently (e.g., within the last 1 month, 3 months, and 12 months) as compared with unmarried women (Table and Figure 1). Data presented further highlight the reality that although marriage is nearly a perfect proxy for experiencing sexual intercourse at some time point (nearly 100% of married women report ever having had sex), it is not a perfect proxy for recent sexual activity. For example, across the 5 regions, on average the percentage of married women reporting sex in the 3 months preceding the interview ranges from 82%–92%. This percentage is lower for sex in the 1 month preceding the interview, where the average reporting ranges from 70%–82% of married women.

Ultimately, limiting the number of women included in mCPR and unmet need estimates to those who report sex in the 1 month preceding the interview provides analysts with sufficiently large numbers of all women and married women for complex analyses (generally more than 5,000 individuals). However, this limit can pose a challenge for analyzing the data of unmarried women, especially in countries and regions

FIGURE 1. Kaplan-Meier Curves for Sexual Recency (Months Before Interview) by Region and Marital Status[a]



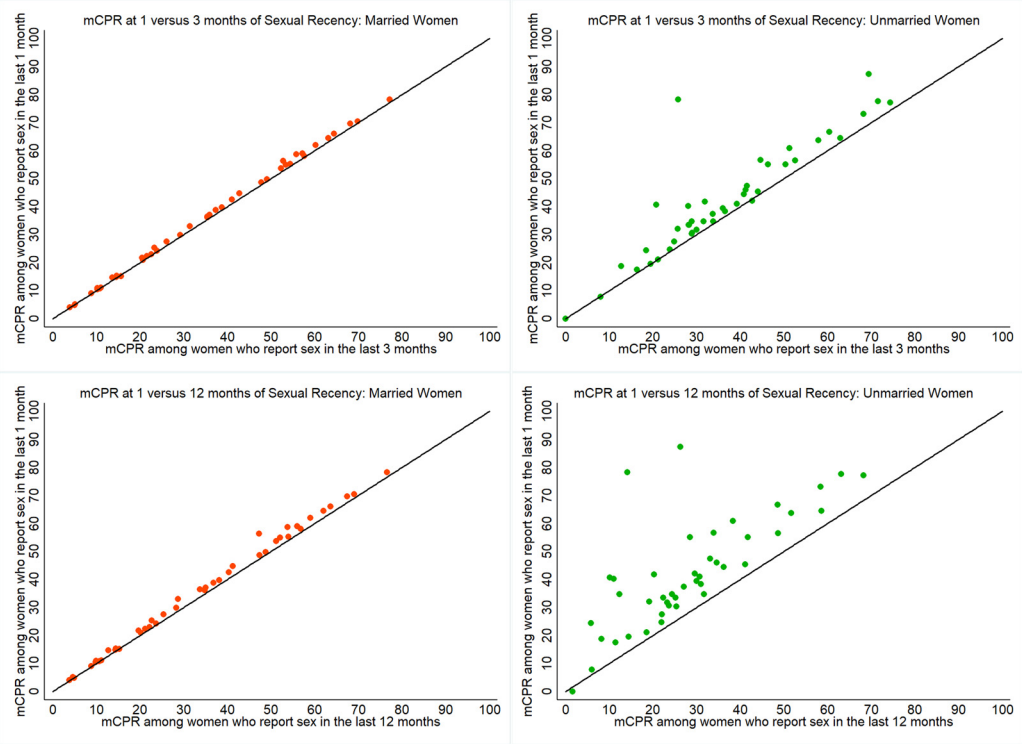
[a] Regional data presented are pooled and weighted at the country level.

where sex outside of marriage is infrequent (or infrequently reported). In the regions of West Asia/Europe and Asia and Pacific, the 1-month cutoff yields fewer than 100 individuals on average. Expanding to those who report sex in the 3 months preceding the interview yields higher numbers in these 2 regions but still less than 250 individuals on average. For national-level estimates of mCPR and unmet need for unmarried women, these sample sizes are generally large enough. For further disaggregation and more complex analyses, these sample sizes are likely too small in many study countries.

How Do mCPR and Unmet Need Estimates Vary by Timing of Last Sex?

Applying these various sexual activity restrictions to the calculation of mCPR for all countries (Figure 2), we observe that for married women, mCPR is virtually the same as measured at the 12-month, 3-month, and 1-month eligibility

FIGURE 2. Modern Contraceptive Prevalence Rate Among Women Who Reported Sex in the Previous 1 Month Compared With Those Who Reported Sex in the Previous 3 Months and 12 Months, by Marital Status for 43 Countries



Abbreviation: mCPR, modern contraceptive prevalence rate.

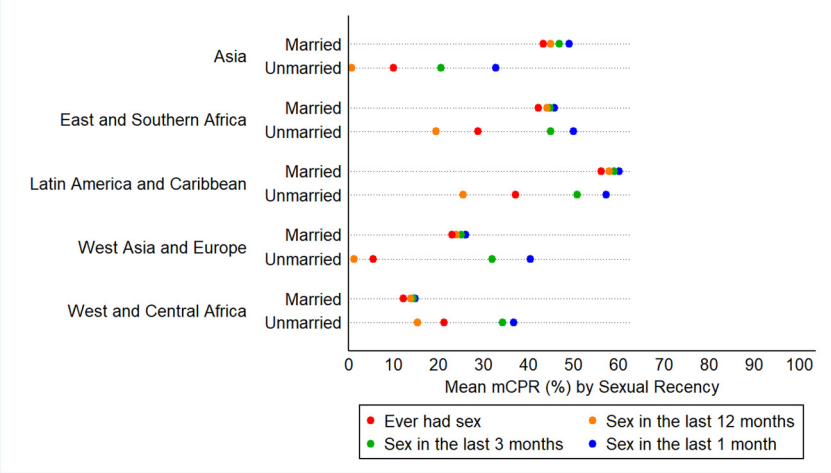
cutoffs. However, for unmarried women, mCPR is systematically lower as measured at the 12-month and 3-month marks as compared with the 1-month cutoff. Indeed, mCPR as measured among unmarried women who reported sex within the 12 months preceding the interview is lower—on average, 14 percentage points lower—across all 43 study countries as compared with mCPR as measured using the 1-month criterion.

These differences are starker when parsed out by regions (Figure 3). Among married women, we see that mCPR estimates based on the sexual recency cutoff of 1 month is virtually indistinguishable from the 3-month cutoff in all regions (varying by 1–2 percentage points). In all regions, the 1-month cutoff results in higher reports of contraceptive use regardless of marital status. Among all regions, married women who have ever been sexually active had the lowest mCPR compared to married women at the other sexual recency cutoffs. Furthermore, unmarried women who report sexual activity in the previous 12 months had the lowest mCPR among both married and unmarried women at the other sexual recency cutoffs (note: mCPR by sexual recency and marital status for all countries is presented in Supplement 1).

Next, we turn our attention to unmet need. The DHS calculation of unmet need among unmarried women applies the inclusion criteria of sex in the 1 month before the interview, which aligns with the DHS mCPR calculation among unmarried women. Because we changed the mCPR calculation among unmarried women to a broader sexual recency time frame, we needed to align our unmet need calculation. Therefore, we recalculated unmet need to expand the inclusion criteria to women who were sexually active within 3 months before the interview and 12 months before the interview versus the current DHS standard of 1 month.

As Figure 4 shows, expanding the inclusion criteria to capture unmarried women with less recent sex results in modest to large increases in

FIGURE 3. Mean Modern Contraceptive Prevalence Rate by Sexual Recency and Marital Status for Each Geographic Region



Abbreviation: mCPR, modern contraceptive prevalence rate

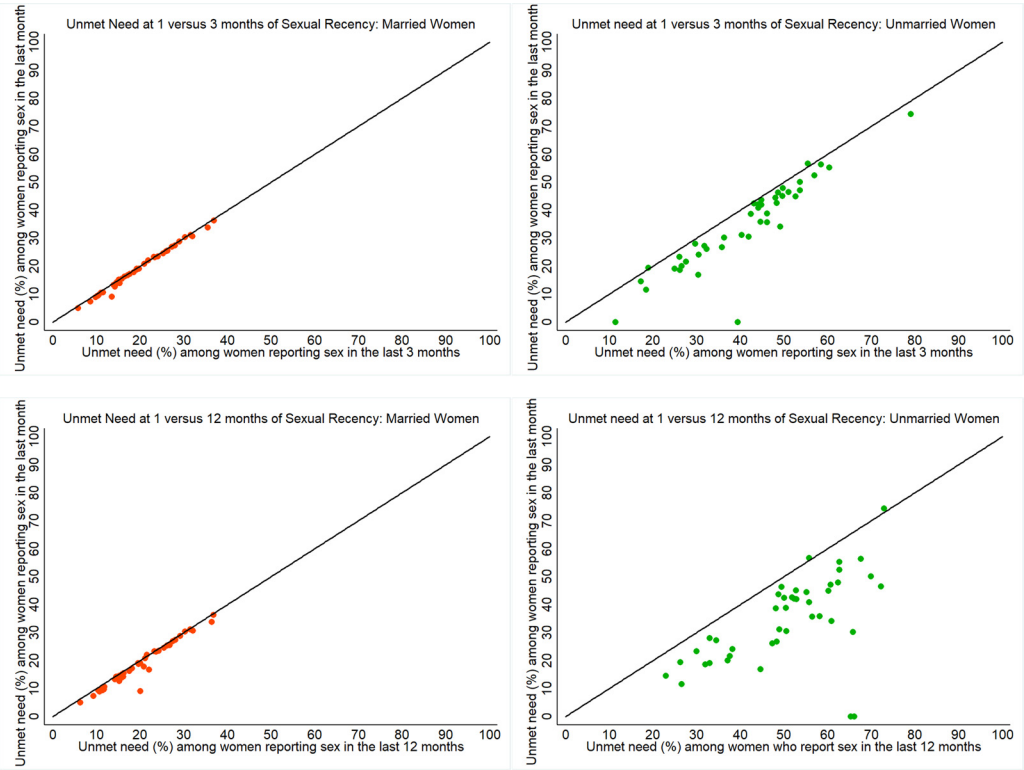
unmet need. Unmet need among unmarried women across our 43 study countries is, on average, 34% (1 month), 41% (3 months), and 51% (12 months). Among married women, the difference is minimal—19% (1 month), 19% (3 months), and 20% (12 months)—as would be expected because mCPR among married women varies little by sexual recency (note: unmet need by sexual recency and marital status for all countries is presented in Supplement 2). Corresponding to the regional differences in mCPR based on the sexual recency cutoffs, the picture for unmet need shows that estimates are lowest when using the 1-month cutoff, and unmet need is highest when using the 12-month cutoff (Figure 5). Unmet need is highest among unmarried women in each of the 5 regions as compared with married women.

IMPLICATIONS

As has been revealed and elaborated in recent scholarship, a substantial proportion of unmarried women have ever had sex.[6,23] Our findings confirm this reality and again highlight the relatively fluid and sporadic nature of unmarried women's sexual lives as compared with married women's (Table). This finding may spur the measurement community to expand the eligibility criteria for mCPR and unmet need calculation to a wider pool of sexually active women (i.e., expand from inclusion criteria of sexual activity within 1 month to 3 months or 3 months to 12 months). Indeed, in so doing, we would be able to have a larger sample size for disaggregation and analysis and perhaps understand contraceptive use dynamics of a wider population of unmarried women.

Our data also reveal that expanding the eligibility range from 1 month to 3 months and beyond yields lower mCPR estimates and higher unmet need estimates. This is to be expected—not because women who had sex less recently are necessarily less likely than others to use contraception or more likely to be in greater need of contraception—but rather because expanding the inclusion criteria based on time since last sex creates a concerning measurement misalignment. Specifically, as previously mentioned, contraceptive use is based on a "current use" measure. This "current" measure aligns well with the experiences of women who had sex more recently. However, in seeking to understand contraceptive use of women whose last sex was 3 or more months ago, a reliable estimate is unlikely to be obtained based on a "current use" question.[11,24] Coital-dependent methods are especially likely to be underreported to a "current use" question among women who are not recently sexually active. To remedy the measurement misalignment, contraceptive use at last sex would need to be examined.

FIGURE 4. Unmet Need Among Women Who Reported Sex in the Previous 1 Month Compared With Those Who Reported Sex in the Previous 3 Months and 12 Months, by Marital Status for 43 Countries

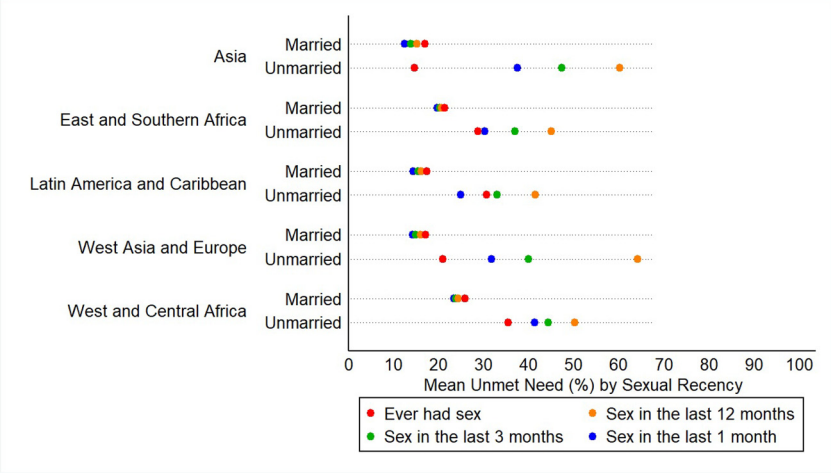


In the DHS, contraceptive use at last sex can be assessed by proxy using the contraceptive calendar. A recent analysis[11] did just that, exploring CPR among 3 groups of women: (1) women who had sex in the last 3 months, by using the contraceptive calendar to assess whether she used contraceptives in the 3 months before the interview; (2) women who had sex in the last 3 months, using the responses to the "current use" questions to define contraceptive use; and (3) women who had sex in the last 1 month, measuring contraceptive use based on responses to "current use." The results showed that CPR was similar in groups 1 and 3, but slightly lower as measured for group 2. Based on these findings, plus the fact that contraceptive calendar data are not as widely available as current use data, the study authors chose to use group 3 as their analytic group. We agree with this assessment.

Because DHS data are cross-sectional, unmarried women who meet the 1-month inclusion criteria are likely to represent all types of sexual activity—frequent, sporadic, and periodic. Therefore, mCPR estimates and unmet need estimates reliably reflect the population-level estimates for this diverse demographic group. Of course, it is possible that there is some degree of seasonality to sexual recency, depending on cultural context, for example, with high levels of male migration. Evidence of seasonality impact on sexual activity is scarce across country contexts. Furthermore, because DHS fieldwork often spans 4 or more months, any potential seasonality impact is diminished. The major downside of this approach, as mentioned earlier, is that sample sizes of unmarried women who met the 1-month inclusion criteria can be quite small in many countries, limiting statistical power for further disaggregation and more complex analysis, especially for family planning among sexually active youth—a key population for international family planning programming. Therefore, we make 2 recommendations:

1. In the short term, we recommend maintaining the DHS method (sexual activity within the previous 1 month) for reporting mCPR and unmet need among unmarried women

FIGURE 5. Mean Unmet Need for Family Planning by Sexual Recency and Marital Status for Each Geographic Region



to more accurately monitor both indicators. For clarity and messaging, we recommend that the wider family planning community adopt this approach for measuring and reporting mCPR and unmet need among unmarried women.

New DHS questions on contraceptive use at last sex can improve calculation of mCPR among unmarried women who had sex less recently.

2. In the longer term, we recommend adding the following 2 questions to the DHS women's questionnaire and other surveys that capture contraceptive use data: (1) "The last time you had sex, did you or your partner do anything or use any method to avoid or delay pregnancy?" (2) [if yes] "What method did you use?" These questions align with questions already asked in the DHS men's questionnaire as well as questions included in the U.S. National Survey of Family Growth.[25] Incorporating these questions would help to overcome the measurement misalignment described herein, thereby allowing for improved calculation of mCPR among unmarried women who had sex less recently and allowing for analysts to use contraceptive use data for a larger number of unmarried women, thus, helping to overcome the sample size issue. For example, this approach would improve our ability to compare mCPR and unmet need estimates by marital status thanks to the greater statistical power it provides. As of editing this article in October 2019, we can report that the DHS Program has incorporated these suggested questions into its DHS-8 core women's questionnaire.[26] We encourage the family planning community to incorporate these new data into contraceptive prevalence and unmet need calculations as the data become available.

Even when both of these recommendations came to fruition, measurement challenges remain. First, in many countries unmarried women face stigma for sexual activity outside of marriage. Such cultural sensitivity may make unmarried women unlikely to report sexual activity, leading to reporting bias.[27,28] Looking more closely at country-specific data (Supplement 1), we observe just how low reported sexual activity rates are among unmarried women in places with large social sanctions on premarital sex. For example, in Nepal, a country with strong norms prohibiting premarital sex, only 4 unmarried women in the entire survey reported sex in the preceding month. Although disclosure of premarital sex is rare, this finding does not necessarily mean that premarital sex itself is rare. Indeed, a body of research is forming that reveals prevalence of premarital sex in Nepal is not uncommon.[29] Some welcome survey innovations are taking place, like using audio computer assisted self interview rather than face-to-face interviews, which

may reduce reporting bias by better capturing responses to sensitive questions like sexual activity and sexual history.[30,31] More can and should be done. For example, question order and wording should be considered to reduce bias. Although DHS asks sexual activity questions much later in the questionnaire—a proven practice for building rapport before arriving at stigmatized issues like sexual behavior[32]—the sexual recency questions come immediately after marital status questions, which may unintentionally invite stigmatized responses from unmarried women in places like Nepal.

Even if we can decrease reporting bias, we still have the challenge of nonresponse and survey representativeness.[31] Specifically, those unmarried women who are sexually active and report sexual activity may have different life experiences than other sexually active unmarried women who are uncomfortable reporting sexual activity. It is possible that unmarried women who report sexual activity are more likely to access and use contraception, which would artificially inflate mCPR among unmarried women and deflate unmet need.

Finally, because sexual activity is more likely to be sporadic among unmarried women compared with married women, it is important to recognize that these women are perhaps more likely to be unprepared (with contraception) for their next sexual encounter and more exposed to the risk of an unintended pregnancy as a result. This is a key challenge that public health programs can and should address. In this vein, it is imperative that the global family planning community commit to having solid, reliable, and comparable data to inform programming to address the family planning needs of unmarried women.

Acknowledgments: We thank the article reviewers and the participants at the 2019 Population Association of America conference for their invaluable comments in improving this paper. We are indebted to the respondents who participated in these DHS surveys.

Competing interests: None declared.

Disclaimer: The views and opinions expressed in this article are the authors'. They do not necessarily represent the views and opinions of the U.S. Agency for International Development.

REFERENCES

1. Brown W, Druce N, Bunting J, et al. Developing the "120 by 20" goal for the Global FP2020 Initiative. *Stud Fam Plann*. 2014; 45(1):73–84. CrossRef. Medline
2. Cahill N, Sonneveldt E, Stover J, et al. Modern contraceptive use, unmet need, and demand satisfied among women of reproductive age who are married or in a union in the focus countries of the Family Planning 2020 initiative: a systematic analysis using the Family Planning Estimation Tool. *Lancet*. 2018;391(10123): 870–882. CrossRef. Medline
3. Darroch JE, Singh S. Trends in contraceptive need and use in developing countries in 2003, 2008, and 2012: an analysis of national surveys. *Lancet*. 2013;381(9879):1756–1762. CrossRef. Medline
4. Sathar Z, Chidambaram VC. *Differentials in Contraceptive Use. WFS Comparative Studies 36*. Voorburg, Netherlands: International Statistical Institute; 1984.
5. *Population Facts: World Marriage Patterns*. 2011/11. New York, NY: United Nations Department of Economic and Social Affairs, Population Division; 2011.
6. MacQuarrie KLD, Mallick L, Allen C. *Sexual and Reproductive Health in Early and Later Adolescence: DHS Data on Youth Age 10–19*. DHS Comparative Reports No. 45. Rockville, MD: ICF; 2017. http://dhsprogram.com/pubs/pdf/CR45/CR45.pdf.
7. Das Gupta M, Engelman R, Levy J, Luchsinger F, Merrick T, Rosen JE. *The Power of 1.8 Billion: Adolescents, Youth and the Transformation of the Future*. State of World Population 2014. New York, NY: UNFPA; 2014. https://www.unfpa.org/sites/default/files/pub-pdf/EN-SWOP14-Report_FINAL-web.pdf. Accessed May 24, 2019.
8. Rimon JG, Tsui AO. Regaining momentum in family planning. *Glob Health Sci Pract*. 2018;6(4):626–628. CrossRef. Medline
9. Kendall MG. *The World Fertility Survey: Current Status and Findings. Population Reports* M(3). Voorburg, Netherlands: International Statistical Institute; 1979.
10. Morris L, Lewis G, Powell D, et al. *Contraceptive Prevalence Surveys: A New Source of Family Planning Data. Population Reports M(5);* 1981.
11. Wang W, Staveteig S, Winter R, Allen C. *Women's marital status, contraceptive use, and unmet need in Sub-Saharan Africa, Latin America, and the Caribbean*. DHS Comparative Report No. 44. Rockville, Maryland: ICF; 2017. http://dhsprogram.com/pubs/pdf/CR44/CR44.pdf. Accessed May 24, 2019.
12. Croft TN, Marshall AMJ, Allen CK. *Guide to DHS Statistics*. Rockville, MD: ICF; 2018.
13. Darroch J. Adding It Up: *Investing in Contraception and Maternal and Newborn Health, 2017 - Estimation Methodology*. New York, NY: Guttmacher Institute; 2018.
14. Hindin MJ, Kalamar AM. Country-specific data on the contraceptive needs of adolescents. *Bull World Health Organ*. 2017;95(3):166. CrossRef. Medline
15. Dasgupta A, Ueffing P, Kantorová V. *Sexual Activity by Marital Status and Age: A Comparative Perspective. United Nations, Department of Economic and Social Affairs, Population Division* Technical Paper No. 11. New York, NY: United Nations; 2017.
16. Fabic MS, Choi Y, Bird S. A systematic review of Demographic and Health Surveys: data availability and utilization for research. *Bull World Health Organ*. 2012;90:604–612.
17. Smith-Greenaway E, Clark S. Variation in the link between parental divorce and children's health disadvantage in low and high divorce settings. *SSM Popul Health*. 2017;3:473–486. CrossRef. Medline
18. Clark S, Brauner-Otto S. Divorce in sub-Saharan Africa: are unions becoming less stable? *Popul Development Review*. 2015;41(4):583–605.
19. Heger Boyle E, King M, Sobek M. *Data Description-TIMESINCESEXFLAG: IPUMS-Demographic and Health Surveys*: Version 7. Minnesota Population Center and ICF International; 2019. https://www.idhsdata.org/idhs-action/variables/TIMESINCESEXFLAG#description_section.
20. Bradley SEK, Croft TN, Fishel JD, Westoff CF. *Revising Unmet Need for Family Planning*. DHS Analytical Studies No. 25. Calverton, MD: ICF; 2012. https://dhsprogram.com/pubs/pdf/AS25/AS25[12June2012].pdf.

21. MacQuarrie KLD. *Marriage and Fertility Dynamics: the Influence of Marriage Age on the Timing of First Birth and Birth Spacing*. DHS Analytical Studies No. 56. Rockville, MD: ICF; 2016. http://dhsprogram.com/pubs/pdf/AS56/AS56.pdf.
22. Bradley SE, Casterline JB. Understanding unmet need: history, theory, and measurement. *Stud Fam Plann*. 2014;45(2):123–150. CrossRef. Medline
23. Kalamar AM, Tunçalp Ö, Hindin MJ. Developing strategies to address contraceptive needs of adolescents: exploring patterns of use among sexually active adolescents in 46 low- and middle-income countries. *Contraception*. 2018;98(1):36–40. CrossRef. Medline
24. Fabic MS, Becker S. Measuring contraceptive prevalence among women who are at risk of pregnancy. *Contraception*. 2017;96(3):183–188. CrossRef. Medline
25. Aiken AR, Wang Y, Higgins J, Trussell J. Similarities and differences in contraceptive use reported by women and men in the National Survey of Family Growth. *Contraception*. 2017;95(4):419–423. CrossRef. Medline
26. Demographic and Health Survey (DHS) Program. DHS Model Questionnaires: Phase 8. Washington, DC: DHS; 2019. https://dhsprogram.com/publications/publication-DHSQ8-DHS-Questionnaires-and-Manuals.cfm. Accessed October 25, 2019.
27. Beguy D, Kabiru CW, Nderu EN, Ngware MW. Inconsistencies in self-reporting of sexual activity among young people in Nairobi, Kenya. *J Adolesc Health*. 2009;45(6):595–601. CrossRef. Medline
28. Kelly CA, Soler-Hampejsek E, Mensch BS, Hewett PC. Social desirability bias in sexual behavior reporting: evidence from an interview mode experiment in rural Malawi. *Int Perspect Sex Reprod Health*. 2013;39(1):14–21. CrossRef. Medline
29. Adhikari R, Tamang J. Premarital sexual behavior among male college students of Kathmandu, Nepal. *BMC Public Health*. 2009;9(1):241. CrossRef. Medline
30. Hewett PC, Mensch BS, Erulkar AS. Consistency in the reporting of sexual behaviour by adolescent girls in Kenya: a comparison of interviewing methods. *Sex Transm Infect*. 2004;80(Suppl 2):ii43–ii48. CrossRef. Medline
31. Langhaug LF, Sherr L, Cowan FM. How to improve the validity of sexual behaviour reporting: systematic review of questionnaire delivery modes in developing countries. *Trop Med Int Health*. 2010;15(3):362–381. CrossRef
32. Fenton KA, Johnson AM, McManus S, Erens B. Measuring sexual behaviour: methodological challenges in survey research. *Sex Transm Infect*. 2001;77(2):84–92. CrossRef. Medline

Peer Reviewed

Received: August 22, 2019; **Accepted:** October 29, 2019

Cite this article as: Short Fabic M, Jadhav A. Standardizing measurement of contraceptive use among unmarried women. *Glob Health Sci Pract*. 2019;7(4):564-574. https://doi.org/10.9745/GHSP-D-19-00298

GLOBAL HEALTH: SCIENCE AND PRACTICE
Dedicated to what works in global health programs
OPEN ACCESS

ORIGINAL ARTICLE

Effects of a Peer-Led Intervention on HIV Care Continuum Outcomes Among Contacts of Children, Adolescents, and Young Adults Living With HIV in Zimbabwe

Talent Tapera,[a] Nicola Willis,[a] Kudakwashe Madzeke,[a] Tanyaradzwa Napei,[a] Mather Mawodzeke,[a] Stanley Chamoko,[a] Abigail Mutsinze,[a] Teddy Zvirawa,[a] Beatrice Dupwa,[b] Aveneni Mangombe,[b] Anesu Chimwaza,[b] Talent M. Makoni,[b] Winnie Mandewo,[c] Mbazi Senkoro,[d] Philip Owiti,[e,f] Jaya Prasad Tripathy,[e,g,h] Ajay M.V. Kumar[e,g,i]

An intervention focused on children, adolescents, and young adults living with HIV using a cadre of dedicated peers—community adolescent treatment supporters—led to improvements along the HIV care cascade among their household contacts and sexual partners.

ABSTRACT

Background: Africaid Zvandiri, in partnership with the Ministry of Health and Child Care (MOHCC) in Zimbabwe, implemented a comprehensive, peer-led program, focused on children, adolescents, and young adults living with HIV aged 0–24 years. The peers, known as community adolescent treatment supporters (CATS), are people living with HIV (PLHIV) aged 18–24 years who are trained and mentored to support their peers throughout the HIV care continuum through support groups, home visits, phone call reminders, and messages. We report the HIV care continuum outcomes (HIV testing uptake, antiretroviral therapy [ART] uptake, retention, and viral suppression) in a cohort of household contacts and sexual partners (aged younger than 25 years) of index children, adolescents, and young adults living with HIV identified by CATS from October 2017 to September 2018 in 24 districts of Zimbabwe.

Methods: This was a retrospective cohort study involving analysis of routine program data, extracted from electronic databases consisting of data on contacts of index PLHIV and ART outcomes. We used April 30, 2019, as the censor date for all analyses.

Results: A total of 15,223 household contacts and sexual partners with unknown HIV status (linked to 9,353 index PLHIV) were identified and referred for HIV testing. Of these, 12,114 (79.6%) were tested and 1,193 (9.8%) were HIV-positive. Of the latter, 1,153 (96.6%) were initiated on ART with 99% starting on the day of diagnosis. Of those on ART, 1,151 (99.8%) were alive on ART at 6 months and 2 (0.2%) died. A total of 1,044 (91%) children, adolescents, and young people living with HIV underwent viral load testing at 6 months or later, of whom 1,037 (99.3%) were virally suppressed (<1000 copies/ml).

Conclusion: These findings add to the global evidence demonstrating the effectiveness of peer-led interventions in children, adolescents, and young adults living with HIV and justify the decision of the MOHCC in Zimbabwe to scale-up the model nationally. Future research should aim to understand the reasons for the gaps in HIV testing and viral load testing using qualitative research.

INTRODUCTION

Since the first HIV/AIDS patient was reported more than 35 years ago, approximately 78 million people globally have become infected with HIV and 35 million have died from AIDS-related illnesses.[1] By the end of 2017, there were 36.9 million people living with HIV (PLHIV), of whom 5.7 million (15%) were children (0–9 years), adolescents (10–19 years), and young adults (20–24 years of age).[1]

Adolescents have an increased tendency for risk-taking behavior, including unsafe sexual practices and substance abuse that increases their vulnerability to acquiring HIV.[2,3] Apart from the risk behaviors, some

[a] Africaid Zvandiri, Harare, Zimbabwe.
[b] Ministry of Health and Child Care, Harare, Zimbabwe.
[c] Elizabeth Glaser Pediatric AIDS Foundation, Harare, Zimbabwe.
[d] National Institute for Medical Research, Muhimbili Centre, Dar es Salaam, Tanzania.
[e] International Union Against Tuberculosis and Lung Disease, Paris, France.
[f] National Tuberculosis, Leprosy and Lung Disease Program, Nairobi, Kenya.
[g] International Union Against Tuberculosis and Lung Disease, South-East Asia Office, New Delhi, India.
[h] All India Institute of Medical Sciences, Nagpur, India.
[i] Yenepoya Medical College, Yenepoya, Mangaluru, India.
Correspondence to Talent Tapera (talent.tapera@gmail.com).

children and adolescents have prenatally acquired HIV given the natural history of untreated HIV infections. Between 2005 and 2012, global HIV-related deaths increased among adolescents by 50% but declined among all PLHIV.[4] Since then, although deaths among adolescents with HIV have been decreasing due to massive scale-up of antiretroviral therapy (ART), the rate of decrease has been slower compared to the rate of decrease among adults.[1]

Compared to adults, adolescents are less likely to access HIV testing and treatment services, remain in care, and achieve viral suppression.

Several studies in sub-Saharan Africa have shown that, compared to adults, adolescents were less likely to access HIV testing and treatment services, remain in care, and achieve viral suppression.[5–11] These results were mainly due to poor prioritization of adolescents in most national HIV plans, inadequate provision of HIV testing and treatment services, delays in diagnosis and treatment, and lack of services to support retention in care.[12] Children also remain a vulnerable and neglected group with issues related to consent, access, acceptability, cultural norms, stigma, and discrimination.[13] HIV testing and ART coverage among children are often lower than among adults.[1]

Zimbabwe has been facing a generalized HIV epidemic with 1.3 million PLHIV (17% aged younger than 25 years) by the end of 2017.[1] A national survey conducted in 2015–2016 showed that only 34% of young adults (15–24 years) knew their HIV status compared to 74% among adults,[14] a figure well below the global UNAIDS 90-90-90 target. Because HIV testing is the gateway to care and treatment, the World Health Organization (WHO) recommends that all the adolescents in settings with generalized epidemics be offered HIV testing and counseling, preferably using community-based approaches including home-based HIV testing and self-testing.[12] The HIV care continuum begins with diagnosis to linkage to ART treatment to being retained in care to viral suppression.

Community adolescent treatment supporters have been trained, mentored, supervised, and incentivized to counsel and support their peers throughout the HIV care continuum.

Since 2002, Africaid Zvandiri, in partnership with the Ministry of Health and Child Care (MOHCC) in Zimbabwe, has been implementing the Zvandiri program. This program is a comprehensive, multicomponent, multidonor funded, differentiated service delivery program focused on the overall development of children, adolescents, and young adults living with HIV that includes HIV care, sexual and reproductive health, mental health, and social protection. At the heart of this program are community adolescent treatment supporters (CATS), adolescents and young adults aged 18–24 years who are living with HIV and who have been trained, mentored, supervised, and incentivized to counsel and offer support to their peers throughout the HIV care continuum, through support groups, home visits, counseling, short message service (SMS), and phone call reminders. Previous program evaluation indicated increased linkage of children, adolescents, and young adults to HIV treatment and retention as well.[15] Encouraged by these positive results, the MOHCC in Zimbabwe scaled up the CATS model in 51 districts (of 63 in Zimbabwe) in 2017.[15]

The role of CATS was expanded in 2016 to include contact investigation, including counseling all the household contacts and sexual partners of index HIV patients, referring them for HIV testing, and linking those diagnosed as HIV-positive to care and support including ART. However, there has not been a systematic assessment of how well this expansion is functioning in routine program settings.

Hence, we undertook an operational research study to assess the effect on HIV care cascade outcomes (HIV testing uptake, ART uptake, ART retention, and viral suppression) in a cohort of household contacts and sexual partners (aged younger than 25 years) of index children, adolescent, and young adults living with HIV identified by CATS from October 2017 to September 2018 in selected districts of Zimbabwe.

METHODS

Study Design

This was a retrospective cohort study involving analysis of secondary data routinely available in records of the Zvandiri program from October 2017 to September 2018. The study was done from January 2019 to May 2019.

Setting

Zimbabwe, a landlocked country situated in southern Africa, has a population of 13.1 million. According to a national survey, 74.2% of PLHIV aged 15–64 years knew their HIV status, 86.8% of those who knew their HIV status received ART, and 86.5% of the latter were virally suppressed.[14]

The Zvandiri Model

The goal of Zvandiri (meaning "as I am") program is to achieve and maintain physical, social, and mental well-being of children, adolescents, and young adults living with HIV. CATS are at the forefront of service delivery in this program. The CATS' major responsibilities in working with

children, adolescents, and young adults living with HIV include: (1) cofacilitating monthly support groups and ART refill groups, (2) conducting home visits, (3) sending SMS reminders and check-ins, (4) counseling, (5) making phone calls, (6) referring and linking them to care, (7) conducting community outreach visits, and (8) cofacilitating caregiver workshops. For younger age groups, the CATS support the children through the parents or caregivers. The CATS are incentivized with a fixed allowance of US$20 per month, bicycles to facilitate home visits and/or reimbursement of bus fare, and monthly airtime allowance for SMS reminders and calls. More details about CATS are summarized in Table 1. The model of care has been described in detail elsewhere and is visually summarized in Figure 1.[15]

Through the CATS' support of index PLHIV, they had an avenue to meet the index cases' contacts. The CATS screened the contacts against the age, biological relations, sexual relationship, and prior testing eligibility parameters. They first sought permission of the contacts during home visits and ensured privacy. Through this screening process, CATS referred sexual partners (aged younger than 25 years) as well as household contacts (aged younger than 25 years) who were biologically related and staying together under the same roof for HIV testing and supported pre- and post-test HIV counseling and disclosure. The contact tracing forms were returned and kept in a lockable cabinet at the health facility. There were 3 options provided to the contacts for HIV testing: facility-based testing, self-testing, and home-based testing. Those who opted for self-testing were provided with self-testing kits and, if found to be reactive, they were referred for confirmatory HIV testing at the nearest health facility. If

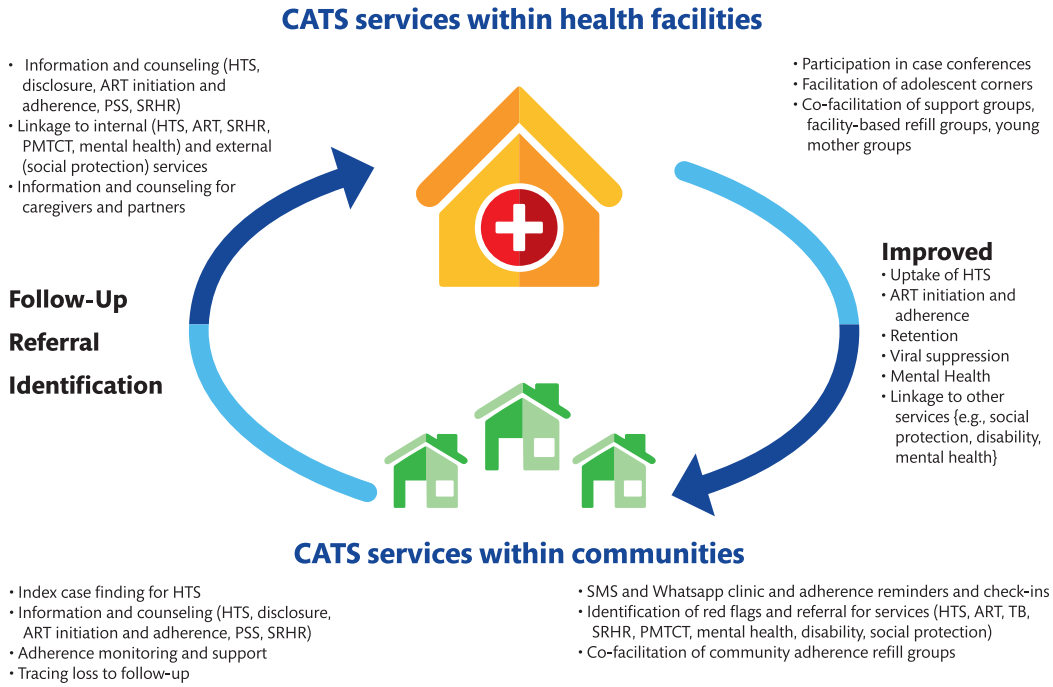
Through the CATS' support of index PLHIV, they had an avenue to meet the index cases' household contacts and sexual partners.

TABLE 1. Description of Community Adolescent Treatment Supporters in the Africaid Zvandiri Program in Zimbabwe

Who are CATS?	Adolescent and young adults living with HIV (18–24 years old) trained and mentored by MOHCC and Africaid as peer counselors
Who appoints CATS?	Health care facility staff identify PLHIV (18–24 years old) with the potential (willing, competent, and motivated) to be CATS and then appoint them in consultation with the authorities in the MOHCC and Africaid. CATS should have completed secondary school and must have consent from their caregivers to enroll.
How many CATS per health facility?	Although the number of CATS per health facility depends on the number of children, adolescents, and young adults living with HIV who need support, the aim is to at least have 1 male and 1 female CATS per health facility. Each CATS should support between 30 and 60 children, adolescents, and young adults living with HIV at any given point in time.
What training do they receive?	All CATS receive 2 weeks of MOHCC-endorsed training on knowledge related to pediatric and adolescent HIV (HIV, ART, adherence support, disclosure, sexual and reproductive health, protection, psychosocial support, and mental health) and skills in counseling and community outreach. The training combines theory and practical components, which includes hands-on mentorship ('shadowing') by senior CATS for a period of time before they are independently able to provide support. Training is participatory and uses case studies and role plays. They also receive technical support from district-based Zvandiri mentors employed by Africaid. This initial training is then followed by continued on-site training and mentorship.
What are their responsibilities?	Cofacilitate monthly support groups and ART refill groups Conduct home visits for counseling, monitoring, and support Send SMS reminders and phone calls for adherence and clinic visits and check-ins Provide counseling in clinic and link to other services as needed Refer children, adolescents, and young adults living with HIV (particularly severe cases) and link to other service providers including OI/ART, mental health, social protection, disability, SRHR, and PMTCT Perform community outreach visits in partnership with other cadres from health and social protection Cofacilitate caregiver workshops
Who supervises and mentors CATS?	A nurse or primary counselor at the clinic supervises CATS with additional supervision and mentorship by the district Zvandiri mentor. A district-level monthly meeting is conducted to mentor and review progress of CATS.
What remuneration and incentives do they receive?	Fixed allowance of US$20 per month Bicycles to facilitate home visits and/or reimbursement of bus fare Monthly airtime allowance for SMS reminders and calls

Abbreviations: ART, antiretroviral therapy; CATS, community adolescent treatment supporter; MOHCC, Ministry of Health and Child Care; OI, opportunistic infection; PLHIV, people living with HIV; PMTCT, prevention of mother-to-child transmission of HIV; SMS, short messaging service; SRHR, sexual and reproductive health and rights.

FIGURE 1. The Zvandiri Model of Care Involving the Community Adolescent Treatment Supporters in Zimbabwe



Abbreviations: ART, antiretroviral therapy; CATS, community adolescent treatment supporters; HTS, HIV testing services; PLHIV, people living with HIV; PMTCT, prevention of mother-to-child transmission of HIV; PSS, psychosocial support; SRHR, sexual and reproductive health and rights; SMS, short message service; TB, tuberculosis.

contacts or partners felt hesitant to come to a health facility, home-based testing was done by roving testers from other organizations. The CATS used a standard MOHCC tool to refer the contacts or partners for HIV testing.

Contacts and partners found to be HIV-negative were linked to HIV prevention services, including voluntary medical male circumcision, family planning, and cervical cancer screening for young women. Contacts confirmed as HIV-positive were registered in the health facilities for initiating ART. Diagnosis and treatment were conducted per national guidelines, which followed a "test and treat" policy, in line with the WHO guidelines.[16,17] In addition to the care contacts and partners received as part of the national program, they were also registered with the Zvandiri program for treatment adherence support and other activities with CATS.

Differentiated Care

Recognizing that not all children, adolescents, and young adults living with HIV required the same level of care, CATS differentiated and tailored services according to the person's clinical, psychological, and social needs and circumstances. The 2 types of support, standard support and enhanced support, are summarized in Table 2.

Briefly, CATS provided standard support to PLHIV who were clinically and psychosocially stable and who regularly attended their clinic visits. CATS provided enhanced support to those who did not regularly attend their clinic visits, were not virologically suppressed, or had special needs (e.g., mental health conditions, pregnant women, and those at risk of or subjected to abuse or neglect). In enhanced support, CATS increased the frequency and intensity of contact with children, adolescents, and young adults with HIV. CATS provided counseling and referred severe cases to the Zvandiri mentors and health facility and other relevant service providers.

Monitoring and Supervision

CATS were attached to the health facilities in their areas. CATS were supervised and supported by the MOHCC staff and district-based Zvandiri mentors employed by Africaid. CATS were expected to

TABLE 2. Components and Levels of Support Provided by CATS to Children, Adolescents, and Young Adults Living With HIV in the Africaid Zvandiri Program, Zimbabwe

	Standard Zvandiri Support	Enhanced Zvandiri Support
Eligibility criteria	• Undetectable viral load or CD4 count >500 cells/ml in the last 6 months • Attended all scheduled clinic visits in the last 3 months • Psychologically stable • Safe	• A detectable viral load or CD4 count <500 cells/ml in the last 6 months • Failed to attend scheduled clinic visits in the last 3 months • Psychological distress • Abuse or neglect • Started ART in the past 3 months • Reported nonadherence • Pregnant
CATS-led interventions	• Monthly home visit • Weekly SMS reminders • Clinic-based counseling • Referrals and linkages, particularly for severe cases, to other service providers	• Home visit every 2 weeks • Daily SMS reminders • Clinic-based counseling • Referrals and linkages, particularly for severe cases, to other service providers
CATS-supported interventions	• Monthly support group • Caregiver workshop	• Monthly support group • Caregiver workshop • Adherence workshop • Community outreach with CHWs/CCCWs

Abbreviations: CATS, community adolescent treatment supporters; CHWs, community health workers; CCCWs, child case care workers.

submit a monthly report in the prescribed format and attend a monthly review meeting chaired by the Zvandiri mentor. This meeting acted as a forum to discuss successes and challenges and decide on the course-corrective actions.

Recording

When CATS visited the index cases and contacts at their homes, they captured the information with respect to HIV testing in a paper-based, structured proforma called an index case form. Clients with previously known HIV status were not listed in the index form. This form was digitized in a Microsoft Excel sheet at the district level. The details of contacts diagnosed as HIV-positive and started on ART were then captured electronically in the ART database, each getting a unique identification code (UIC).

Study Population and Period

The study population comprised household contacts aged younger than 25 years (defined as people who were biologically related and staying together under the same roof) and sexual partners (aged younger than 25 years) of the index PLHIV, identified by CATS from October 2017 to September 2018 in 24 selected districts of Zimbabwe. These districts were purposively chosen based on availability of data (of contacts) in the electronic form. People with a previously known HIV status were not included.

Data Collection

Data variables included age, sex, HIV testing (yes/no), HIV test result, ART start (yes/no), ART outcomes (alive and on ART/death/loss to follow-up) at 6 and 12 months of care, viral load test, and test result. Data were sourced from 2 databases: (1) contact database, containing the details of the contacts and HIV testing, and (2) ART database, containing the details of ART outcomes and viral load testing. For each HIV-positive client identified in the contact database, we searched the ART database using the UIC, name, and sex to assess if the client has been initiated on ART. This merged database was used for all analysis. The censor date for assessing all the

outcomes was April 30, 2019, when the data were downloaded from the ART database.

Data Analysis

Data analysis was done using EpiData (v2.2.2.187, EpiData Association, Odense, Denmark) and STATA (version 12, Texas, USA) software. HIV testing, HIV positivity, ART initiation, retention in care, and viral suppression (defined as <1000 copies per ml) were summarized using frequencies and proportions. The operational definitions of the outcomes for children, adolescents, and young adults living with HIV were as follows: (1) **Alive and on ART**: those who did not miss their clinical visits and documented to have received care at the time of assessment; (2) **Death**: those who died at any time during the course of treatment; and (3) **Loss to follow-up**: those who were not seen at the ART center for 90 days or more after their scheduled appointment and who could not be contacted successfully.

Ethical Approval

Ethics approval was obtained from the Medical Research Council of Zimbabwe (MRCZ/E/225) and the Ethics Advisory Group of International Union Against Tuberculosis and Lung Disease, Paris, France (EAG 61/18). Because the study involved a review of existing records without direct interaction with human participants, the need for informed consent was waived by the ethics committees.

RESULTS

A total of 15,223 household contacts and sexual partners linked to 9,353 index PLHIV (1.6 contacts per index case) were identified by CATS during the study period. There were 278 CATS (approximately 34 index PLHIV per CATS on average) in the study areas. Of the 278 CATS, 155 (55%) were women and 123 were men (45%).

Of the contacts identified and referred for HIV testing by CATS, nearly 80% were tested.

The demographic characteristics of the contacts are shown in Table 3. A majority of the contacts were women (57%). Most of the contacts were either children (57%) or siblings (41%) of the index cases, and the remaining (2%) were sexual partners.

TABLE 3. Demographic Characteristics of the Household Contacts and Sexual Partners of Index People Living With HIV Identified by Community Adolescent Treatment Supporters in 24 Districts of Zimbabwe, October 2017–September 2018 (N=15,523)

Characteristics	No. (%)
Age (years)	
0–4	2495 (16.4)
5–9	2814(18.5)
10–14	3170 (20.8)
15–19	3830 (25.2)
20–24	2914 (19.1)
Gender	
Male	6597 (43.3)
Female	8626 (56.7)
Type of Contact	
Sibling	6229 (40.9)
Sexual partner	290 (1.9)
Children	8704 (57.2)
Province	
Bulawayo	1567 (10.3)
Harare	685 (4.5)
Manicaland	4489 (29.5)
Mashonaland Central	1076 (7.1)
Mashonaland East	179 (1.2)
Masvingo	1362 (8.9)
Matabeleland South	1871 (12.3)
Midlands	3994 (26.2)

HIV Testing

All the 15,223 contacts and sexual partners were referred for HIV testing by the CATS. Of these, 12,114 (79.6%) were tested for HIV, and 1,193 (9.8%) were found to have HIV infection (Table 4). Most (98%) of the contacts were tested for HIV at a health facility and the rest underwent self-testing or home-based testing. In multivariable analysis, age and province were independently associated with not testing for HIV (Table 5). Not testing was significantly higher among younger age groups (i.e., those aged 0–4 years and 5–9 years) when compared to the 20–24 years age group.

ART Linkage

Of the 1,193 HIV-positive contacts identified, 1,153 (96.6%) were initiated on ART. Of the latter, 1,144 (99.2%) were initiated on the same day

TABLE 4. HIV Care Cascade Among Household Contacts and Sexual Partners of Index PLHIV Identified by Community Adolescent Treatment Supporters in 24 Districts of Zimbabwe, October 2017–September 2018 (N=15,223)

	No. (%)
HIV Testing Outcome	
Contacts Referred for HIV Testing	15,223 (100.0)
Contacts tested for HIV	12,114 (79.6)
Contacts tested who were HIV-positive	1,193 (9.8)
HIV-positive contacts who initiated ART	1,153 (96.6)
ART Outcomes	
3 months	
Alive on ART	1,153 (100.0)
6 months	
Alive on ART	1,151 (99.8)
Died	2 (0.2)
12 Months (n=569[a])	
Alive on ART	566 (99.5)
Died	2 (0.4)
Lost to follow-up	1 (0.2)
Viral suppression (<1000 copies/ml)	1,037[b] (99.3)

Abbreviations: ART, antiretroviral therapy; PLHIV, people living with HIV.
[a] Number of contacts eligible for 12-month assessment; People whose duration between ART start date and censor date was less than 12 months were considered not eligible for assessment.
[b] Among 1,044 contacts who had a viral load test at 6 months or later after starting ART.

of testing, and the remaining started within a week.

ART Outcomes

Of 1,153 contacts who started ART, 1,151 (99.8%) were alive ART at 6 months, and 2 (0.2%) had died. At the time of censoring, many clients had not completed 12 months since starting ART and were not eligible for the 12-month assessment. Of the 569 who were eligible for assessment, 566 (99.5%) were alive and on ART, 2 (0.4%) people had died, and 1 (0.1%) was declared lost to follow-up.

Viral Suppression

Of 1,153 people who started ART, 1,044 (91%) had a viral load test conducted at 6 months or later and of them, 1,037 (99.3%) were found to be virally suppressed.

DISCUSSION

This study adds to the growing body of evidence demonstrating the effectiveness of targeted, peer-led, differentiated care delivery models in improving the HIV care outcomes among children, adolescents, and young adults living with PLHIV. The CATS model in Zimbabwe is one such successful intervention. We found high rates of HIV testing (80%), ART uptake (97%), retention in care (99%), viral load testing (90%), and viral suppression (99%) in a large cohort of household contacts and sexual partners of index PLHIV under the care of CATS in Zimbabwe. These are excellent outcomes by any standard and better than those reported nationally in Zimbabwe.[14] A population-based study in Zimbabwe reported that among PLHIV aged 15–24 years, only 50.4% knew their HIV status, 83.7% self-reported receiving ART, and 85.4% were virally suppressed.[14] In this study, 83.7% of the 15–24 year olds who were referred for testing managed to get tested.

This peer-led care model resulted in high rates of HIV testing, ART uptake, retention in care, viral load testing, and viral suppression.

TABLE 5. Factors Associated With Not Testing for HIV Among Household Contacts and Sexual Partners of Index PLHIV Identified by Community Adolescent Treatment Supporters in 24 Districts of Zimbabwe, October 2017–September 2018

Characteristic	Total Referred for HIV Testing	Not Tested for HIV	Crude RR (95% CI[a])	Adjusted RR[b] (95% CI[a])
	No.	No. (%)		
Total	15223	3109 (20.4)		
Age (years)				
0–4	2495	864 (34.6)	13.63 (10.82–17.18)	12.72 (10.01–16.17)
5–9	2814	775 (27.5)	10.84 (8.59–13.68)	10.41 (8.19–13.23)
10–14	3170	867 (27.4)	10.77 (8.54–13.58)	10.20 (8.03–12.96)
15–19	3830	529 (13.8)	5.43 (4.28–6.90)	5.21 (4.08–6.65)
20–24	2914	74 (2.5)	Ref	Ref
Gender				
Male	6597	1456 (22.1)	1.15 (1.08–1.22)	1.05 (0.98–1.13)
Female	8626	1653 (19.2)	Ref	Ref
Type of contact				
Sibling	6229	1289 (20.7)	Ref	Ref
Sexual partner	290	12 (4.1)	0.19 (0.11–0.34)	0.51 (0.29–0.91)
Children	8704	1808 (20.8)	1.00 (0.94–1.06)	1.04 (0.96–1.12)
Province				
Masvingo	1362	55 (4.0)	Ref	Ref
Bulawayo	1567	365 (23.3)	5.76 (4.38–7.58)	7.27 (5.47–9.67)
Harare	685	2 (0.3)	0.07 (0.01–0.29)	0.08 (0.02–0.36)
Manicaland	4489	1321 (29.4)	7.28 (5.60–9.47)	7.36 (5.62–9.65)
Mashonaland Central	1076	75 (7.0)	1.72 (1.23–2.42)	1.73 (1.22–2.45)
Mashonaland East	179	9 (5.0)	1.24 (0.62–2.47)	1.34 (0.66–2.72)
Matabeleland South	1871	812 (43.4)	10.74 (8.25–13.99)	11.84 (9.00–15.57)
Midlands	3994	470 (11.8)	2.91 (2.21–3.82)	3.32 (2.51–4.40)

Abbreviations: CI, confidence interval; PLHIV, people living with HIV; Ref, reference group; RR, risk ratio.
[a] Factors with confidence intervals not including 1 were statistically significant ($P<.05$).
[b] Adjusted for age, sex, province, and type of contact.

The major gap in the HIV treatment cascade was with HIV testing.

The good outcomes reported in this study may be attributed to the following reasons: (1) a structured and well-defined CATS model implemented in close collaboration with the MOHCC; (2) systematic training of CATS using a standard training curriculum reinforced by continuous on-the-job mentorship by Africaid mentors; (3) close follow-up of the clients by CATS using home visits, SMS reminders, phone calls, and support groups, which concur with the findings of a cohort study in Tanzania that suggested that providing additional psychosocial support to PLHIV receiving ART can reduce loss to follow-up[18]; (4) supportive supervision and monitoring of CATS by the government health care providers; (5) incentives that included a fixed allowance of US$20 per month, airtime allowance for making phone calls and sending SMS reminders to clients, and a bicycle to make home visits; and (6) nonfinancial incentives and motivators such as recognizing the best-performing CATS during the monthly meetings and providing them an opportunity to travel and mentor other CATS in the neighboring districts.

The major gap in the HIV treatment cascade was at the level of HIV testing, where about 20% of the contacts were not tested for HIV.

Although we do not know the exact reasons for this gap, we speculate some reasons based on the program experience and multivariable analysis showing the associations of age and province with HIV testing.

HIV testing coverage was low among younger age groups, with the lowest coverage in under-5 children. This may be related to legal barriers, requiring consent of the parent or caregiver.[19] The CATS ensured that for children requiring caregiver consent for testing, the caregivers were engaged. Although several steps have been taken in Zimbabwe to address this issue including educating the caregivers through workshops, the gap remains. Current efforts target only caregivers of children, adolescents, and young adults living with HIV after HIV diagnosis.[15] These efforts need to be expanded, and caregivers of all contacts should be educated. There may be other reasons that include cultural norms dictating that young infants should not be taken outside the house except for receipt of vaccines (not always colocated with early infant diagnosis), transportation barriers, a child having to miss school, and the persistent stigma of HIV/AIDS and fear about discrimination.[13] The CATS screen out children who may have gone through early infant diagnosis for eligibility of HIV testing referral.

HIV testing was lowest in Matabeleland South province. We hypothesize that this may be because this is a border province and many contacts identified initially may have moved across the border to South Africa before HIV testing. This finding needs further investigation. The uptake of self-testing and home-based testing was low in our study. Strengthening these may have potential to fill the gaps in HIV testing. Also, not testing was lowest among the sexual partners (4.1%) compared to 20.7% for siblings and 20.8% for parents. Thus, the program seems to be successful in reaching out to the sexual partners of HIV-positive adolescents and young people.

A strength of the study was that we had a large sample covering 24 districts (of the total of 63) in Zimbabwe, making the findings more generalizable to other areas implementing the CATS model.

Program Implications

There are a couple of program implications. First, a system should be instituted to routinely capture the reasons for not testing, non-initiation of ART, and other adverse program outcomes. This will enable periodic assessments of the reasons for the gaps in the HIV care cascade and course correction. Second, we identified some inconsistencies in recording including duplicate records. Data on dates of HIV testing and viral load testing were missing, which would have enabled us to assess the delays involved in the process. These need to be corrected, and measures of data quality assurance and quality control should be put in place.

Limitations

There were some limitations to the study. First, we did not have a control group in our study, which would have enabled a direct head-to-head comparison. Second, we also did not assess the reasons for the gaps in the cascade of care. Third, we did not collect the data on costs, which would have enabled a cost-effectiveness analysis. These knowledge gaps will be addressed by a cluster randomized trial that is underway.[20] Fourth, we relied on routine program data, which may have had recording errors. Fifth, there was no information on other sociodemographic and clinical factors associated with not testing. So, there could be some bias due to these unexplained confounders.

CONCLUSION

In conclusion, we found high levels of HIV testing and care outcomes among a cohort of household contacts and sexual partners of index PLHIV who received care by the CATS in Zimbabwe. Contacts of index cases is an additional component of the CATS program that can reach and benefit children, adolescents, and young adults who are HIV-positive and out of care, expanding the potential impact of CATS. Future assessments should focus on exploring the reasons for the gaps in the HIV cascade using qualitative research.

Acknowledgments: This research was conducted through the Structured Operational Research and Training Initiative (SORT IT), a global partnership led by the Special Programme for Research and Training in Tropical Diseases at the World Health Organization (WHO/TDR). The training model is based on a course developed jointly by the International Union Against Tuberculosis and Lung Disease (The Union) and Medécins sans Frontières (MSF). The specific SORT IT program that resulted in this publication was implemented by the Centre for Operational Research, The Union, Paris, France. Mentorship and the coordination/facilitation of this particular SORT IT workshop was provided through the Centre for Operational Research, The Union, Paris, France; the Department of Tuberculosis and HIV, The Union, Paris, France; The Union, Zimbabwe Office; The Union, South East Asia Office; University of Washington, School of Public Health, Department of Global Health, Seattle, Washington, USA; National Institute for Medical Research, Muhimbili Centre, Dar es Salaam, Tanzania; and AIDS & TB Department, Ministry of Health & Child Care, Harare, Zimbabwe.

Funding: Funding for program implementation was provided by the U.S. President's Emergency Plan for AIDS Relief through the U.S.

Agency for International Development, Centers for Disease Control and Prevention, International Training and Education Center for Health, and the United Kingdom Department for International Development.

Competing interests: None declared.

REFERENCES

1. Joint United Nations Programme on HIV/AIDS (UNAIDS). *Global HIV & AIDS Statistics - 2018 fact sheet*. UNAIDS website. http://www.unaids.org/en/resources/fact-sheet. Published 2018. Accessed August 9, 2018.
2. Ssewanyana D, Mwangala PN, van Baar A, Newton CR, Abubakar A. Health risk behaviour among adolescents living with HIV in sub-Saharan Africa: a systematic review and meta-analysis. *BioMed Res Int*. 2018;2018:7375831. CrossRef. Medline
3. Toska E, Pantelic M, Meinck F, Keck K, Haghighat R, Cluver L. Sex in the shadow of HIV: a systematic review of prevalence, risk factors, and interventions to reduce sexual risk-taking among HIV-positive adolescents and youth in sub-Saharan Africa. *PLoS One*. 2017;12(6):e0178106. CrossRef. Medline
4. Joint United Nations Programme on HIV/AIDS (UNAIDS). *Global Report UNAIDS Report on the Global AIDS Epidemic 2013*. http://files.unaids.org/en/media/unaids/contentassets/documents/epidemiology/2013/gr2013/UNAIDS_Global_Report_2013_en.pdf. Accessed August 9, 2018.
5. Nachega JB, Hislop M, Nguyen H, et al. Antiretroviral therapy adherence, virologic and immunologic outcomes in adolescents compared with adults in southern Africa. *J Acquir Immune Defic Syndr*. 2009;51(1):65–71. CrossRef. Medline
6. Hudelson C, Cluver L. Factors associated with adherence to antiretroviral therapy among adolescents living with HIV/AIDS in low- and middle-income countries: a systematic review. *AIDS Care*. 2015;27(7):805–816. CrossRef. Medline
7. Kim SH, Gerver SM, Fidler S, Ward H. Adherence to antiretroviral therapy in adolescents living with HIV: systematic review and meta-analysis. *AIDS*. 2014;28(13):1945–1956. CrossRef. Medline
8. Adejumo OA, Malee KM, Ryscavage P, Hunter SJ, Taiwo BO. Contemporary issues on the epidemiology and antiretroviral adherence of HIV-infected adolescents in sub-Saharan Africa: a narrative review. *J Int AIDS Soc*. 2015;18(1):20049. CrossRef. Medline
9. Bobat R, Archary M, Lawler M. An update on the HIV treatment cascade in children and adolescents. *Curr Opin HIV AIDS*. 2015;10(6):411–419. CrossRef. Medline
10. Govindasamy D, Ferrand RA, Wilmore SMS, et al. Uptake and yield of HIV testing and counselling among children and adolescents in sub-Saharan Africa: a systematic review. *J Int AIDS Soc*. 2015;18(1):20182. CrossRef. Medline
11. Callahan T, Modi S, Swanson J, Ng'eno B, Broyles LN. Pregnant adolescents living with HIV: what we know, what we need to know, where we need to go. *J Int AIDS Soc*. 2017;20(1):21858. CrossRef. Medline
12. World Health Organization (WHO). *HIV and Adolescents: Guidance for HIV Testing and Counselling and Care for Adolescents Living with HIV*. Geneva, Switzerland: WHO; 2013.
13. Kilewo C, Massawe A, Lyamuya E, et al. HIV counseling and testing of pregnant women in sub-Saharan Africa: experiences from a study on prevention of mother-to-child HIV-1 transmission in Dar es Salaam, Tanzania. *J Acquir Immune Defic Syndr*. 2001;28(5):458–462. CrossRef. Medline
14. Ministry of Health and Child Care (MOHCC). Zimbabwe. *Zimbabwe Population-Based HIV Impact Assessment (ZIMPHIA) 2015-16: First Report*. Harare, Zimbabwe: MOHCC; 2017. https://phia.icap.columbia.edu/wp-content/uploads/2017/11/ZIMPHIA_First_Report_FINAL.pdf. Accessed August 9, 2018.
15. Willis N, Napei T, Armstrong A, et al. Zvandiri—bringing a differentiated service delivery program to scale for children, adolescents, and young people in Zimbabwe. *J Acquir Immune Defic Syndr*. 2018;78(Suppl 2):S115–S123. CrossRef. Medline
16. World Health Organization (WHO). Consolidated Guidelines on the Use of Antiretroviral Drugs for *Treating and Preventing HIV Infection: Recommendations for a Public Health Approach, 2nd ed*. Geneva, Switzerland:WHO; 2016. http://www.who.int/hiv/pub/arv/arv-2016/en/.
17. AIDS and TB Programme, Ministry of Health and Child Care (MOHCC). *Operational and Service Delivery Manual for the Prevention, Care and Treatment of HIV in Zimbabwe*. Harare, Zimbabwe: MOHCC; 2017. http://ophid.org/treat-all-toolkit/MOHCC Guidelines/MOHCC Zimbabwe OSDM 2017.pdf.
18. Siril HN, Kaaya SF, Smith Fawzi MK, et al. CLINICAL outcomes and loss to follow-up among people living with HIV participating in the NAMWEZA intervention in Dar es Salaam, Tanzania: a prospective cohort study. *AIDS Res Ther*. 2017;14(1):18. CrossRef. Medline
19. Eba PM, Lim H. Reviewing independent access to HIV testing, counselling and treatment for adolescents in HIV-specific laws in sub-Saharan Africa: implications for the HIV response. *J Int AIDS Soc*. 2017;20(1):21456. CrossRef. Medline
20. Mavhu W, Willis N, Mufuka J, et al. Evaluating a multi-component, community-based program to improve adherence and retention in care among adolescents living with HIV in Zimbabwe: study protocol for a cluster randomized controlled trial. *Trials*. 2017;18(1):478. CrossRef. Medline

Peer Reviewed

Received: June 6, 2019; **Accepted:** October 15, 2019; **First Published Online:** December 16, 2019

Cite this article as: Tapera T, Willis N, Madzeke K, et al. Effects of a peer-led intervention on HIV care continuum outcomes among contacts of children, adolescents, and young adults living with HIV in Zimbabwe. *Glob Health Sci Pract*. 2019;7(4):575-584. https://doi.org/10.9745/GHSP-D-19-00210

GLOBAL HEALTH: SCIENCE AND PRACTICE
Dedicated to what works in global health programs
OPEN ACCESS

ORIGINAL ARTICLE

Using Data to Keep Vaccines Cold in Kenya: Remote Temperature Monitoring With Data Review Teams for Vaccine Management

Mercy Lutukai,[a] Elizabeth A. Bunde,[b] Benjamin Hatch,[c] Zoya Mohamed,[a] Shahrzad Yavari,[d] Ernest Some,[e] Amos Chweya,[a] Caroline Kania,[f] Jesse C. Ross,[d] Carmit Keddem,[b] Yasmin Chandani[a]

Using technology to make data visible to stakeholders and giving those stakeholders a framework for analyzing that data for decision making improves cold chain management of vaccines in Kenya.

ABSTRACT

Background: Global vaccination coverage rates have remained around 85% for the past several years. Increasing immunization coverage rates requires an effective cold chain to maintain vaccine potency. Remote temperature monitoring (RTM) technology for vaccine refrigerators has shown promise for improving the ability of supply systems to maintain optimal temperature conditions to ensure potent vaccines reach the end users.

Methods: A pilot study of RTM technology and data use teams was implemented in 36 study sites in Kenya. Data were collected at baseline and endline points over a 3-month baseline and 7-month implementation period. Data included 44 qualitative interviews, process logs, meeting minutes from data use team meetings, and quantitative temperature and power data from the RTM devices.

Results: The ability of cold chain equipment to maintain World Health Organization-recommended temperatures in study sites improved markedly between the baseline and implementation periods, resulting in an improvement in total time spent in the correct range from 83.9% in the baseline period to 90.9% in the intervention period and an improvement in time spent in the too cold range from 6.5% to 1.5%. Friedman tests revealed that differences in time spent in the correct range and time spent in the too cold range during the course of the study were statistically significant ($P<.001$ and $P=.04$, respectively). Qualitative and quantitative data suggest that this improvement was due to a combination of improved responsiveness to temperature excursions at the facility level, resulting from SMS alarms for temperature excursion periods, and improved ability at the management level to recognize and address recurring problems.

Conclusion: The combination of using RTM technology with a structured data review process by a management team is a promising approach for improving cold chain outcomes. Future research examining the added value of each of the technological and behavioral components separately is needed.

INTRODUCTION

Globally, in 2016 more than 5.6 million children died before their fifth birthday, mostly from preventable causes.[1] Immunization has been recognized as one of the most successful public health interventions, but global vaccination rates have remained stagnant at 85% for the past several years.[2–4] The World Health Organization (WHO) estimates that improving vaccine coverage rates could prevent an additional 1.5 million deaths per year.[4] To achieve the high immunization coverage rates needed, effective cold chain management for maintaining vaccine potency is required.[2] Remote temperature monitoring (RTM) technology allows for real-time vaccine cold chain equipment (CCE) temperature monitoring and also provides an avenue for CCE data visibility and use. This enables better monitoring of CCE performance. However, few studies exist about how to integrate it into public health supply chains in a way that ensures data are used for action and decisions and to ensure investments are cost-effective.

WHO standards define an adverse heat event as occurring when vaccines experience a temperature above 8°C for a period of 10 hours or more. An adverse freezing event occurs when vaccines experience a temperature below −0.5°C for a period of 1 hour or more, reflecting the greater general sensitivity of vaccines to freezing

[a] inSupply Health, Nairobi, Kenya.
[b] JSI, Boston, MA, USA.
[c] JSI Research & Training Institute, Inc., Boston, MA, USA.
[d] Nexleaf Analytics, Los Angeles, CA, USA.
[e] Ministry of Health, Republic of Kenya, Nairobi, Kenya.
[f] Nexleaf Analytics, Nairobi, Kenya.
Correspondence to Benjamin Hatch (benjamin_hatch@jsi.com).

than to heat events.[5] WHO guidelines recommend storing vaccines between 2°C and 8°C at all levels of the cold chain because exposure to heat or cold outside that range can adversely affect the immunological properties of the vaccines and thus reduce their potency.[6] Administering compromised vaccines will not provide the intended immune response to protect the vaccinated client and that, in turn, can prevent countries from effectively reaching their coverage targets.[6] A number of studies have shown that exposure to temperature extremes within vaccine supply chains is relatively common in both developed and developing countries. As much as 37% of vaccines are exposed to temperatures below the recommended range in lower-income countries, making this a critical issue to address.[7–13]

Despite the importance of ensuring appropriate temperature ranges for the storage of vaccines throughout the cold chain, relatively few studies have examined the use and effectiveness of temperature monitoring practices or existing studies have found monitoring practices to be substandard, commonly resulting in exposure to temperature extremes.[13] For example, a study of North West Region, Cameroon, found that only 76% of health facilities examined had a functioning thermometer for their vaccine storage unit, and of those, 20% were experiencing abnormal temperatures at the time of data collection.[14] In addition, lack of information about what is happening in the vaccine cold chain at the intermediate and facility levels, particularly about the state of functionality of vaccine storage equipment and the exposure of vaccines to temperature extremes at the last mile of the distribution network, is common among many cold chains in developing countries.[15] As discussed below, Kenya's vaccine supply chain suffers from many of these same problems.

Using remote temperature monitoring to record and report temperature data is a promising innovation to increase access to this information.

The National Vaccine and Immunization Programme (NVIP) manages Kenya's vaccine cold chain. At the central level, NVIP stores the vaccines that UNICEF procures for the Expanded Programme on Immunization (EPI). On a quarterly basis, the central level distributes vaccines to 8 regional stores, and subcounty level vaccine stores collect the vaccines from the regional stores. On a monthly basis, more than 6,900 immunizing health facilities collect their vaccines from the subcounty hospitals.

To gauge the status, strengths, and weaknesses of vaccine management in the NIVP, in November 2013, Kenya used the WHO-UNICEF Effective Vaccine Management Tool to conduct an effective vaccine management assessment.[16] The tool assesses each level of the immunization supply chain and makes recommendations to address areas of weakness. The 2013 effective vaccine mangement assessment found that the NVIP's efforts to increase immunization coverage and prevent disease were significantly hampered by compromised vaccine potency resulting from a lack of CCE preventive maintenance and timely repair, outdated equipment inventories, a shortage of spare parts, and poor temperature monitoring by health care workers. These factors, combined with vaccine stock availability issues, hindered Kenya's efforts to increase immunization coverage and prevent vaccine-preventable diseases.

In addition to the problems noted in the assessment, CCE temperatures at facilities and subcounty stores are manually tracked and recorded using the Fridge-tag 2 (FT2), a continuous temperature monitoring logger. The FT2 has a number of documented problems related to users' lack of knowledge on its use, how to read and interpret FT2 readings, and how to initiate action in response to temperature excursions.[16] Lack of temperature data visibility at different levels of the health system further compounds the challenges in Kenya since temperature data at the facility level uses a paper-based recordkeeping system.

RTM technology for real-time recording and reporting of refrigerator temperature data is a promising innovation to increase access to this information. However, as with any technology, RTM technology alone is not enough to ensure optimal outcomes in maintaining ideal temperature ranges. For example, a study in Laos found that although remote reporting of temperature data was successful, additional training was required to enable data managers to effectively use the data and translate it into effective decision making, highlighting the importance of addressing health worker behavior in addition to technical solutions.[15] Similarly, Comes et al. identified real-time temperature monitoring as a promising technology for transforming the performance of cold chains but concluded that there are major gaps in the research into how information gets used by decision makers in the field to support improvements in the functioning of the cold chain.[17]

We hypothesize that cold chain managers in Kenya do not currently have sufficient data to monitor the performance of their cold chain equipment, are not effectively using the data they have, and are not empowered to effectively escalate issues to higher-level decision makers, who lack appropriate visibility into cold chain performance. To address these gaps, we designed our

study to assess the effectiveness of combining the use of RTM technology and a problem solving approach with a data use team that included members from multiple administrative levels, including both ground-level implementers and higher-level decision makers. The RTM devices were deployed in service delivery sites to facilitate access to real-time temperature and power availability data of vaccine refrigerators. Service delivery sites in similar geographic areas were overseen by a data use team, which used systematic data use and problem solving approaches for addressing temperature excursions and cold chain equipment malfunctioning. The combined technology and behavioral approach provided insight into how to protect vaccine potency through improved cold chain management practices and equipment performance in Kenya.

METHODS

Intervention Description

The study intervention included 2 components to address both equipment and behavior issues: an RTM data collection system and a structured team approach to data review.

First, RTM devices were installed in 59 refrigerators in 36 health facilities and subcounty vaccine stores located in the intervention area. The RTM system consisted of 2 major parts, the hardware and the dashboard. The hardware was a global system for mobile communication (GSM)—that is, connected to a cellular network—with temperature sensor probe(s) that were placed inside a vaccine refrigerator, with the main body of the device positioned nearby, usually mounted on a wall. The system uploaded temperature and grid power availability data to a server using cellular networks. Every 10 minutes, the system collected and sent continuous temperature data to an online dashboard. When temperature excursions occurred, the systems sent SMS text messages to key personnel and emitted audible alarms. Long battery life (up to 3 days) helped ensure continuous operation in the event of a power outage.

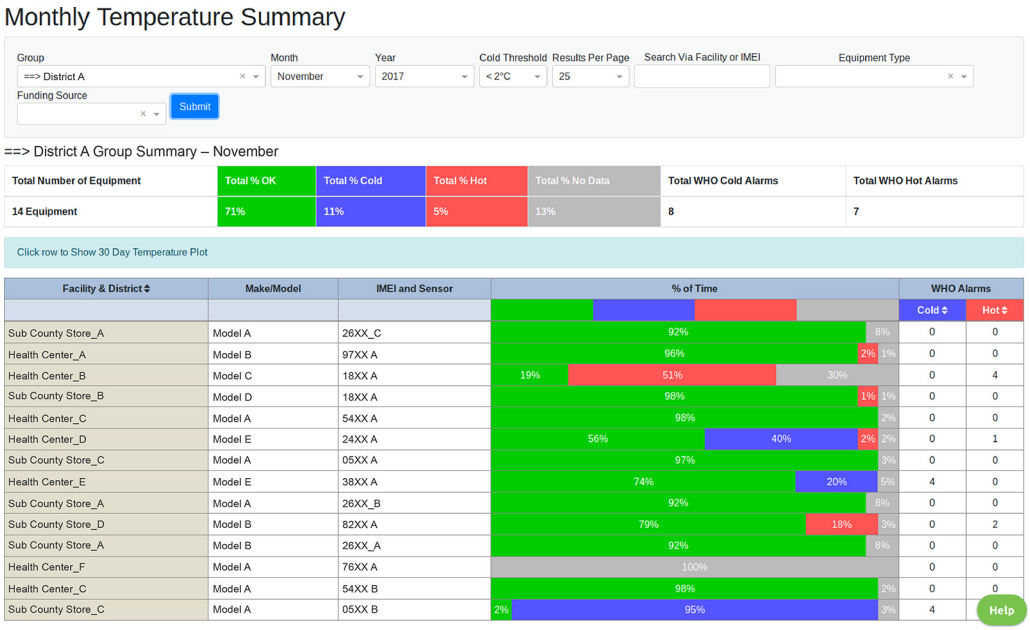
The second part of the RTM system, the dashboard, organized and displayed the collected data through various visualizations and analytics to inform decision making for technicians and managers. Standard visualizations showed each refrigerator's performance as the percentage of time each temperature probe (and refrigerator) measured in each of 3 temperature bands (below, within, and above the WHO recommended temperature range), as well as the number of alarms recorded by the devices each month. Cold chain handlers previously had country-specific standard operating procedures (SOPs) that detailed how to record data from the standard FT2 devices. We provided them with updated SOPs that differed in that they explained how to respond to temperature excursion alarms from the RTM system, maintain cold chain equipment, and escalate unresolved cold chain issues to the county and national levels to be addressed when appropriate. These SOPs were posted near the vaccine refrigerator to help clinic personnel respond to RTM alerts effectively. Facility personnel were also requested to complete process logs to describe actions taken upon encountering alarms.

The second component of the intervention focused on improving behavior of cold chain personnel and improving data use through data use teams. This was a structured approach to team data review modeled after the logistics control tower approach used by many private sector logistics firms. The approach emphasized a discrete team charged with overseeing the performance of the supply chain, selecting indicators to measure performance, using data to track those indicators on a regular basis, and making decisions to address any problems or performance deficiencies identified. These data use teams were comprised of health facility nurses, subcounty and county biomedical engineer technicians, vaccine depot nurses, EPI logisticians, and health records information officers. A set of key performance indicators, derived from the data produced by the RTM devices, were jointly selected with NVIP. These indicators were monitored on a monthly basis and were disaggregated by county, CCE model, and type of facility for analysis. The key indicators selected included number of excursions (high or low temperatures outside of the acceptable range of 2°C to 8°C), percentage uptime (percentage of total time a CCE spent in the range of 2°C to 8°C), and field holdover time (the average amount of time a vaccine fridge in the field maintained safe temperatures after a power outage).

The study intervention assessed equipment and behavior issues in cold chain management.

Having a team with members from multiple disciplines enabled the team to collectively gain a more complete picture of the performance of the supply chain, instead of each member focusing only on the indicators most familiar to them while neglecting others. During the monthly data use team meetings, the teams reviewed performance against key indicators outlined in a jointly established performance plan; identified performance problems; performed root-cause analysis of such

Having a team with members from multiple disciplines provided a more complete picture of the supply chain performance.



Monthly Temperature Summary

Group: ==> District A | Month: November | Year: 2017 | Cold Threshold: < 2°C | Results Per Page: 25 | Search Via Facility or IMEI | Equipment Type | Funding Source | Submit

==> District A Group Summary – November

Total Number of Equipment	Total % OK	Total % Cold	Total % Hot	Total % No Data	Total WHO Cold Alarms	Total WHO Hot Alarms
14 Equipment	71%	11%	5%	13%	8	7

Click row to Show 30 Day Temperature Plot

Facility & District	Make/Model	IMEI and Sensor	% of Time	WHO Alarms	
				Cold	Hot
Sub County Store_A	Model A	26XX_C	92% 8%	0	0
Health Center_A	Model B	97XX A	96% 2% 1%	0	0
Health Center_B	Model C	18XX A	19% 51% 30%	0	4
Sub County Store_B	Model D	18XX A	98% 1% 1%	0	0
Health Center_C	Model A	54XX A	98% 2%	0	0
Health Center_D	Model E	24XX A	56% 40% 2% 2%	0	1
Sub County Store_C	Model A	05XX A	97% 3%	0	0
Health Center_E	Model E	38XX A	74% 20% 5%	4	0
Sub County Store_A	Model A	26XX_B	92% 8%	0	0
Sub County Store_D	Model B	82XX A	79% 18% 3%	0	2
Sub County Store_A	Model B	26XX_A	92% 8%	0	0
Health Center_F	Model A	76XX A	100%	0	0
Health Center_C	Model A	54XX B	98% 2%	0	0
Sub County Store_C	Model A	05XX B	2% 95% 3%	4	0

Help

An anonymized screenshot of the remote temperature monitoring dashboard showing equipment performance statistics. © 2019 NexLeaf Analytics

problems and brainstormed solutions; and developed or updated the team's action plan to address these problems. Recognition of achievements and good performance also served to motivate members to continue striving for performance outside of team meetings. RTM data and the RTM performance dashboard featured as a key component of the data review process during team meetings, though additional program indicators such as vaccine coverage were also tracked.

Study Sites

The study was implemented in Isiolo, Kajiado, and Nairobi counties. These counties were selected because they were identified as priority counties by the Ministry of Health under their health systems strengthening work stream, were participant counties in the Reach Every District, Reach Every Child strategy, and had already established data use teams for vaccine supply chain management. Further, these counties were representative of the different climatic conditions and geographies of other counties in Kenya. Across the 3 counties, 36 study sites were selected to include 18 subcounty vaccine stores and 18 service delivery points with high volumes of vaccine administration. With limited resources to implement the RTM devices, these sites were specifically chosen due to their high volume of vaccine throughput, both to maximize the potential effect of the intervention and because such high-volume sites would be the first targets in any eventual wider-scale adoption of the intervention. These sites represented a small minority of total sites in each county, including 6 of the 46 immunizing sites in Isiolo County (13%), 10 of the 174 sites in Kajiado County (6%), and 20 of the 444 sites in Nairobi County (5%); in total, the 36 sites represented 0.5% of the approximately 7,020 immunizing sites in Kenya.

Study Design

The study used a nonrandomized, pre- and posttest intervention design to determine the efficacy of a combined approach of RTM system implementation for continuous temperature monitoring at the facility and store level with structured data review for action processes by health personnel at multiple levels.

Institutional Review Board approval for this study was not sought, because program leadership, including the principal investigator and supervisors of the implementing team, determined that these activities constituted quality improvement rather than human subjects research. This determination was supported by the focus on a standard programmatic process that would be improved by the RTM concept; the involvement

of internal program staff rather than outside evaluators; and the primary goal of informing operational and strategic decision making. Nonetheless, approval of all activities was obtained from the Kenya Ministry of Health, and informed consent was obtained and documented from all subjects interviewed during the pre- and postintervention periods.

Data Collection and Analysis

Baseline and endline interviews (Supplement) were conducted with EPI personnel, facility in-charges, public health nurses, vaccine depot managers, medical engineering technicians, and health records and information officers. These personnel were involved in either CCE use and management or CCE data performance monitoring at the facility, subcounty, and county levels.

Baseline data were gathered from July to September 2017, including qualitative interviews with 13 total health personnel at study sites. During this time, RTM devices were installed, and the devices recorded and transmitted temperature and power data to the RTM dashboard. During the baseline period, the devices were not configured to send SMS alarms, and the data use team members and facility managers were not provided access to the online RTM dashboard. Qualitative interviews were conducted with 13 EPI staff and cold chain personnel at various levels at each study site to gather information on staff's knowledge of vaccines and current cold chain management practices. While health workers were trained during the baseline period in using the process logs and new SOPs, they continued to follow protocols outlined in the existing SOPs for monitoring refrigerator temperatures using the standard FT2 loggers and paper charts.

The intervention period ran from October 2017 to April 2018. At the start of the intervention period, the RTM system was activated to begin sending audible and SMS alarms to health personnel for temperature excursions and power outages. Key managers at all levels of the system were provided access to data on the online RTM dashboards. Data use teams were also provided with intensive technical support from October to December 2017 to reinforce the routine structured data use team process, including interpreting and reviewing key cold chain indicators via the RTM dashboard and reviewing key supply chain metrics from Kenya's District Health Information System 2 (DHIS 2) system already being used by data use teams.

At the end of the intervention period, qualitative interviews were conducted with 31 total health personnel at study sites. If possible, the individuals interviewed during baseline were also interviewed at endline. Questions included similar knowledge and practice questions as at baseline to provide a comparative understanding of knowledge and perceptions before and after the intervention period. Additional questions on their experiences with the RTM devices, experiences with the data use teams, and ongoing challenges in their vaccine management work were included to retrospectively capture RTM and process-related information.

The study documented changes in key metrics related to vaccine refrigerators' functioning and performance of health worker and cold chain technicians/teams in responding to temperature excursions and maintenance needs. Data used in the analysis came from the baseline and endline qualitative interviews, temperature and power data recorded by the RTM system, written process logs at each site describing alarms and corrective actions taken, and minutes from data use team meetings.

Key themes examined by the qualitative interviews included knowledge about the effect of heating and freezing on vaccines; knowledge and perceptions of the causes of heating and freezing events; recognition of damaged vaccines and the current procedures in managing heat/freeze events and affected stock; and perceived barriers and problems respondents currently face in managing and responding to temperature excursions.

This study measured both the average time spent in excessive temperature zones as well as the number of such heat and freeze excursions (signaled via alarms) that occurred in each refrigerator, based on the temperature and time-series data available in the dashboard. Excursions outside the appropriate temperature range can be indicative of a number of conditions and can help pinpoint appropriate corrective action and redirection of resources. For example, excessive heat alarms may be indicative of frequent power disruptions without appropriate back-up sources of power, or excessive cold alarms may be due to an improperly set thermostat. Excessive cold or heat alarms may also indicate older or poorly functioning equipment that require enhanced preventive maintenance to ensure optimal functionality.

Key indicators from the quantitative temperature data included the percentage of time spent in each temperature band and the numbers of hot and cold alarms calculated by the dashboard

according to WHO-defined temperature excursions of a 10-hour period spent hotter than 8°C for a heat alarm or 1 hour spent colder than −0.5°C for a freeze alarm. A Friedman test was run to determine if there were differences in uptime performance during the 10-month study. Pairwise comparisons were performed (SPSS Statistics, 2018) with a Bonferoni correction for multiple comparisons. This nonparametric test was considered most appropriate because our data did not meet critical assumptions around normality, lack of outliers, and sphericity required for validity with a repeated measures ANOVA.

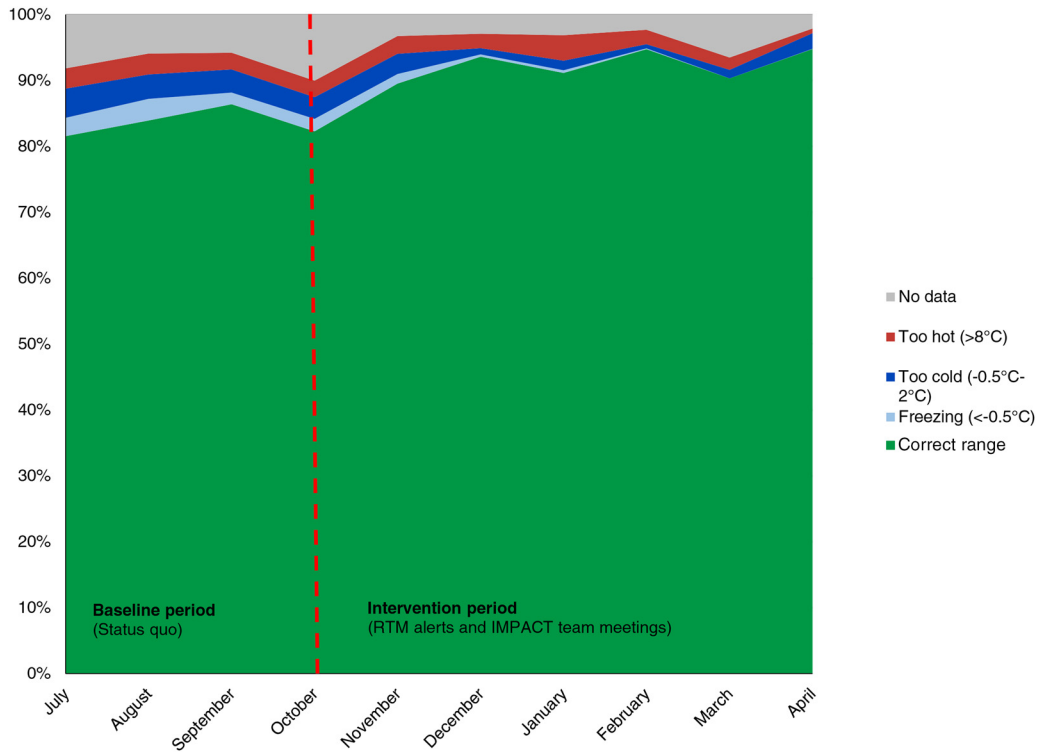
For data analysis, some data from 9 refrigerators were removed due to faulty sensors or refrigerators not in use so as not to skew results. For example, at some sites where a refrigerator was malfunctioning, health staff discontinued use of the refrigerator by unplugging it and removing vaccine supplies to a different refrigerator or facility but did not report the fridge use discontinuation to the study team. However, the RTM device was often left on the refrigerator and continued to transmit data. Where this could be documented, the data from these devices were removed from analysis for the period of time that the refrigerator was not in use. Additionally, 1 of the 18 health facilities selected for the study was removed from analysis entirely as the installation team was unable to find cellular network coverage at the site to enable the RTM device to transmit data.

RESULTS

Temperature Data

Temperature monitoring data revealed a steady improvement in the time that vaccine refrigerators spent in the correct temperature range throughout the implementation period ("uptime"). During the baseline phase, all refrigerators were within the correct temperature range for an average of 83.9% of total time, compared

FIGURE 1. Mean Percentage of Time Spent in Temperature Bands for Vaccine Refrigerators by Month, Baseline (July–September 2017) vs. Implementation (October 2017–April 2018)



with 90.9% of total time during the intervention phase, as seen in Figure 1. According to the Friedman test, uptime performance was statistically significantly different during the different months of the study, $\chi^2(9) = 168.412$, $P<.001$. Post hoc analysis revealed that later months of the study were generally not statistically different from each other and earlier months were not statistically different from each other and but earlier and later months were statistically significantly different from each other.

Although there was improvement in the time spent in the ange of being too hot, it was less dramatic, decreasing from 2.9% of total time during baseline to 2.3% during the intervention phase, and a Friedman test revealed this change to not be statistically significant. The most notable result was a sharp decrease in the time refrigerators spent in the combined ranges of being too cold and freezing, from 6.5% during the 3-month baseline phase to 1.5% during the final 3 months of implementation. A Friedman test for these combined ranges revealed that the time vaccines were exposed to cold temperatures was statistically significantly different during the different months of the study, $\chi^2(9)=17.663$, $P=.04$. This represents a huge reduction in vaccine exposure to inappropriately cold and possibly freezing temperatures.

Across all counties between the baseline and implementation period, the time spent in excessive temperature zones decreased. There was a very slight decrease in number of heat alarms, from an average of 16.3 alarms per month during baseline to an average of 15.3 alarms per month during the implementation period. However, across all counties between the baseline and implementation period there was a marked decrease in freeze alarms, from an average of 65.3 alarms per month during baseline to an average of 21.1 alarms per month during the implementation period. Figure 2 summarizes the average number of monthly freeze alarms by county. However, Friedman tests for the differences in alarms showed that these differences do not meet the threshold of statistical significance ($P=.15$ for cold alarms, $P=.10$ for hot alarms).

One phenomenon that contributed to the reduction in temperature alarms was the identification and repair of thermostats in fridges identified as problematic during data use team meetings. This can be illustrated in Dagoretti subcounty store. Figure 3 shows the temperature oscillation patterns inside a refrigerator from October through December 2017. In October, the temperature oscillated between 2°C and below freezing, causing 56 freeze alarms, with 90% of total time during that month spent below the 2°C threshold. Data use teams and maintenance logs showed that the problem was identified and fixed during the month of November, after which the same oscillating pattern was observed but the oscillations all happened within the appropriate temperature range. This led to the refrigerator spending 97% of total time in December within the correct temperature zone (with 3% of time at unknown temperature), with no freezing events.

There was a huge reduction in vaccine exposure to inappropriately cold and possibly freezing temperatures.

FIGURE 2. Average Number of Freeze Alarms per Month, by County, Baseline (July–September 2017) vs. Implementation (October 2017–April 2018)

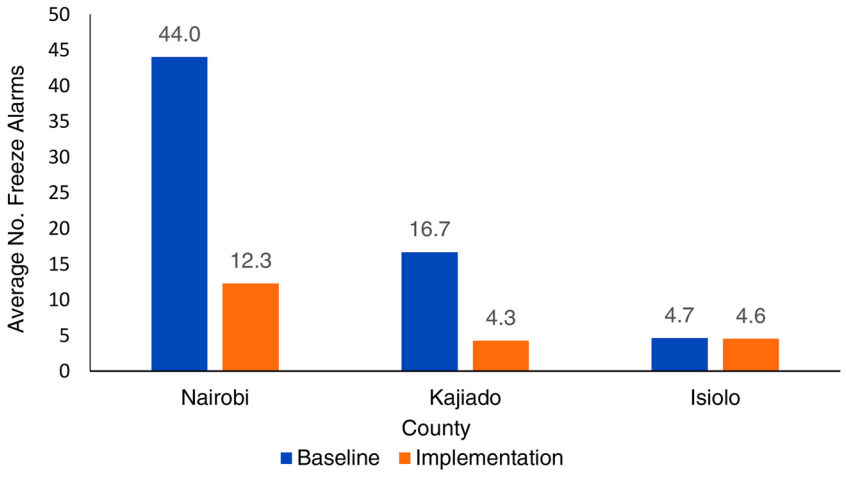
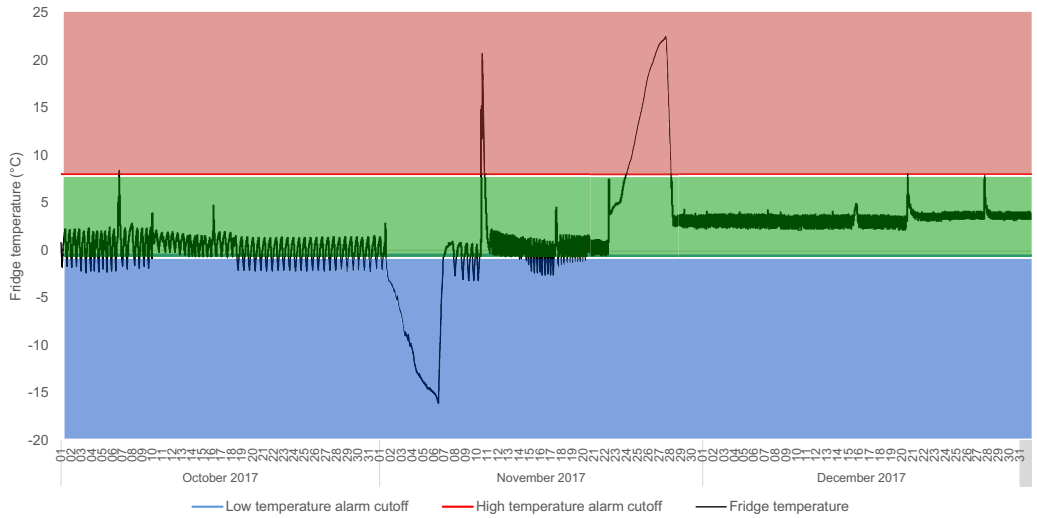


FIGURE 3. Temperature Oscillation of a Refrigerator with a Malfunctioning Thermostat, Dagoretti Subcounty Store, October 2017 to December 2017



Behaviors and Practices

Most interviewees at endline noted personnel errors and suboptimal management practices as contributing factors to temperature excursions. Incorrect vaccine storage or packing procedures, such as placing freeze-sensitive antigens in a freezer or in the wrong compartment of a fridge, were commonplace. Other cited vaccine management errors included forgetting to defrost the refrigerator and frequent and unnecessary opening of the refrigerator. Such behavioral issues were also noted at baseline when respondents indicated that providers were often "careless" in excessively opening the refrigerator, thus increasing stress on the equipment and potentially risking heat exposure. These practices were often the result of personnel using the refrigerators for personal reasons, such as to cool a soda on a hot day. Additionally, respondents noted that problems in facilities with poorly performing equipment were exacerbated by these poor management practices.

Interviewees agreed that having the monitoring system and alarms helped raise awareness and increase responsiveness to temperature excursions.

Beyond these behavior patterns, respondents generally agreed that having the RTM system and receiving alarms helped them to be more aware of and responsive to temperature excursions. In terms of facilitating workflow, most survey respondents agreed that having the RTM devices made their jobs much easier, and almost all respondents described the "alarm-to-action" provided by the RTM devices as having been beneficial to their ability to monitor the state of the vaccine refrigerators and identify and respond to temperature excursions in a timely manner. For example, one facility in-charge noted:

> *I think it has changed my work in a way that it's very easy because even if I'm not in the facility, the moment I get the alarm, I just communicate to one of the staff who is on duty to go and check if the temperatures are going up and if the power is off. It is very easy to manage even if you're far away and you don't have to be within the facility.*

Data Use

The RTM system alarms played an important role in enhancing the use of RTM data. However, data use teams played an equally important role in improving cold chain equipment outcomes because the team structure helped to address system or management issues. Interviewees noted that the processes and data used were very helpful, particularly in identifying recurring challenges (e.g., problematic thermostats) and common malpractices regarding cold chain maintenance (e.g., infrequent refrigerator defrosting that affected cold chain performance). As an example, poorly functioning equipment at 1 site had long been a source of concern for Ministry of Health supervisors. Review of RTM data at a data use team meeting revealed that these concerns were justified because this site was a standout poor performer. Having the

relevant decision makers discuss this issue together resulted in the refrigerator finally being replaced.

Several respondents also noted that the data use team meetings provided a continual educational forum where common issues and awareness of preventive maintenance practices could be raised and shared. It created a space for all personnel to discuss the data, get insight into the problem by asking questions, and try to find solutions to consistent cold chain failures. As a county logistician stated:

> *My opinion is that it's where we meet and really share the data. When you don't share the data, it's like we are in the darkness. The meeting really helps us to see the data, and see whether we're performing or not and especially on the antigens because in that dashboard for vaccines, we are able to see that the vaccines I am having are they enough, I'm out of stock or overstocked. So all those things we share in the [data use team] meetings and if all the subcounties are together, it helps and we are able to fix a problem.*

Finally, many respondents also embraced the RTM dashboard as a useful technology that helped improve CCE data integrity and accuracy. Through the dashboard, CCE medical engineer technicians and biomedical technicians were easily able to continuously monitor their CCE performance, identify mistakes in their records, and make appropriate corrections. Facility-level nurses also valued the dashboard as revealing the extent of vulnerability in the cold chain. As a nurse stated, "It is an eye opener how vaccines have been exposed to cold and heat excursions."

DISCUSSION

Although well-performing equipment is important for cold chain management, other aspects also play a key role in ensuring vaccine potency, including trained personnel, effective and efficient SOPs, and management practices that link trends and priorities in cold chain management with supply chain management and service provision.

The study results indicate that the combined intervention components of the RTM system with the structured data use team approach to data use and problem solving had a direct, positive impact on cold chain management outcomes by ensuring timely action as triggered by SMS alarms and addressing recurring challenges to improve systems overall. Within these results, we outline 3 main conclusions.

Technology Benefits Can Be Enhanced When Matched With Effective Problem Solving and Decision Making Processes

Qualitative interviews clearly indicated that the SMS alarms provided by the RTM system to health staff served as a cue to timely action and fostered greater awareness of the performance of cold chain equipment. In triangulating with the quantitative data, the improvements in performance appeared to be due to the adoption of improved management practices resulting from data use teams' enhanced ability to identify recurring problems and take action to address them.

We see this most clearly in the dramatic change in the number of freeze alarms and time spent in the cold temperature range for a few sites. Our results showed that not all refrigerators that spent a significant amount of time below −0.5°C also reported multiple freeze alarms. RTM devices were configured to be more sensitive to cold excursions, as exposure to freezing temperatures has a more immediate negative impact on potency than excessive heat. However, even among freezing alarms, each alarm can indicate a different message about the status of the cold chain. For example, a prolonged freezing period would produce only 1 alarm indicating the equipment stays in the freezing range, whereas multiple relatively short periods of freezing would produce multiple alarms. Although each of these problems may be caused by an incorrectly set or malfunctioning thermostat within the refrigerator, the multiple alarms may indicate a small adjustment is needed rather than a complete thermostat replacement.

The differences in freeze alarms was not obvious to facility-level staff who simply receive the alarms from the system, but data use teams were able to detect this pattern by looking at the time-series data and alarm records used during their data review. Often multiple freeze alarms from a given sensor in a particular month were due to the refrigerator's thermostat being set too low. This resulted in the refrigerator's compressor automatically turning on to cool the unit, which would push the refrigerator's temperature below that −0.5°C mark, thus triggering an alarm after an hour. When the compressor turned off again, the refrigerator would warm up again to just above the −0.5°C mark, resetting the alarm and turning the compressor on again, thus reinitiating the cycle. Consequently, the refrigerators trigger multiple alarms as the temperature cycles back and forth across the −0.5°C mark.

Trends over time, such as these oscillating temperature patterns, were more easily observed through the time-series data available. Combined with the problem solving process during data use team meetings, these problematic thermostats were identified and actions planned and undertaken to adjust or replace them. The reduction in number of freeze alarms is therefore indicative of the ability of the dashboard data and data use teams to identify problematic refrigerators and prioritize them for repair or adjustment.

The composition of the data use teams, which included not just implementing facility-level staff but also engineers and logisticians from the subcounty and county levels with more power to set maintenance agendas, allowed these recurring issues to be escalated and prioritized beyond the facility level where the issues were occurring. Though limitations in spare part availability and travel budgets at times hindered the ability of the health system to address some of these issues, once issues were flagged, they could often be addressed more efficiently, leading to quicker improvements in performance after implementation of the intervention. This was particularly effective in the case of refrigerators that were experiencing regular freezing events due to an incorrectly set thermostat. Once flagged and addressed, a sharp decrease in the time vaccines were exposed to the combined temperature ranges of being too cold and freezing was observed, as outlined above.

Ability to Triangulate Many Data Sources Is More Likely to Facilitate Holistic Problem Solving

Systems are dynamic and complex with many interrelated issues. Thus, instilling a data use culture for effective cold chain management is facilitated when data are used to address a variety of related system bottlenecks rather than focusing too narrowly on a single issue. Our qualitative results showed that the strategy of expanding existing data use teams that were already looking at coverage and supply chain metrics and adding RTM related metrics enabled teams to triangulate their data, indicators, and results to problem solve more effectively and inform the range of decisions and actions to be taken.

The data use team member feedback during the implementation process also showed that having data available from a variety of sources for multiple indicators encouraged the teams to operationalize the data use concepts. This more robust understanding of the overall situation served as a way to unify the perspectives of the different members of the data use teams since they could each explain their own data in a way that related to overall program performance. Generally, the data use team meetings were described as "very helpful," a forum to "create teamwork," and an enabling platform to "discuss issues and evaluate performance." Some respondents also noted the data use team meetings "helped improve reporting indicators such as vaccine coverage," "discuss commodity shortages and wastage rates," and "monitor temperature excursions" to improve conditions.

For example, initially the service providers (nurses) on the team were most interested in service provision statistics provided by existing supply chain data dashboards, since they considered this their primary performance metric. Meetings would begin with reviewing that data first and interpreting data from the RTM dashboard within the context of how it would affect the service provision statistics. They noted that the RTM system performance indicators would have been less interesting and meaningful without that context. In contrast, the technicians on the team were more interested in RTM dashboard data. They are unaffected by service provision data but the RTM data directly affected their decisions and actions. Having both types of data and indicators available, along with members with both perspectives, helped to emphasize the system linkages between the indicators and the drivers of performance challenges. As a result, the team was able to more effectively understand and resolve issues.

Having multiple categories of data also helps the data use teams develop a deeper understanding of the performance of the supply chain. When only a single data source is available and performance is meeting targets, little triangulation of data occurs. Consequently, the data use team is unlikely to take any further action and the value of the forum becomes less evident. Reviewing the different types of data required problem solving and collaboration across team members with different perspectives who are used to working in individual silos, thus promoting shared understanding and shared accountability. The RTM and logistics dashboards capture indicators across these different components and enabled triangulation of data, holistic problem solving, and action planning around cross-cutting programmatic issues.

Staff-Level Knowledge and Practices Are Key for Long-Term Systemic Change

The qualitative data demonstrated that there are still basic skill and knowledge gaps among both

cold chain personnel and EPI staff that were present at both the baseline and endline points. The data use team intervention did not include a training component to address skill and knowledge gaps specifically, which likely affected the effectiveness of the intervention since some health workers may not have been equipped to take corrective actions agreed upon in meetings. For example, at both baseline and endline there was still confusion among several of the respondents about which vaccines were heat or cold sensitive, or the time before the vaccine is considered damaged once the excursion had happened. These knowledge gaps can have important implications for potency. Even if RTM data identified an excursion, without health worker capacity to perform a vaccine vial monitor staging or shake test to determine if the temperature excursion has damaged the vaccine, or even knowledge on basic procedures for managing vaccines within the fridge to minimize damage, an investment in RTM devices will be unlikely to achieve the benefits the technology offers. Many respondents also expressed worry that many cold chain personnel had not practiced these skills or had not had refresher training since their original vaccine management training.

The study results revealed the need for knowledge and skills development (new or refresher) to be included as follow-up actions from data use team meetings and incorporated into the data use team meetings themselves or for a larger initiative to improve knowledge and skills to be implemented as a complement to RTM system implementation and use.

Limitations

Several limitations affected the results of the study. Limited budgetary allocations to support the continuation of the data use team meetings throughout the implementation period affected the regularity of the meetings and the ability of health personnel to maximize the use of data in their decision making. Althought the data use team meetings were to transition to health facility in-charges meetings starting in January 2018 and supported by local budgetary allocations, not all counties were able to sustain the transport and meeting costs (standard per diem for participants and conference room rental). In some of the Nairobi subcounties, the costs were mitigated by partner resources available to support the meetings, but this support was not consistent across all subcounties, limiting the number of monthly meetings held throughout the intervention period. This affected the ability of the intervention to ascertain the full impact of the behavioral elements supported by the data use team.

The political election period, which started in July 2017 and continued through October 2017, was generally disruptive to the health system and specifically to the study implementation. Several subcounties were unable to hold data use team meetings in October 2017 and health workers were not present at facilities for prolonged periods to monitor or repair cold chain equipment in the event of breakdowns or to address repair needs. Unfortunately, this disruption was coincident with our baseline data collection and may reasonably be expected to have impacted responsiveness to temperature excursion alarms. We were not able to control for this effect in our analysis. The elections also resulted in staffing changes across all 3 counties and at all levels including county directors for health, facility, and EPI staff. This affected the composition of the data use teams. New data use team members were not trained in the approach and were inexperienced in data use team processes, which affected the momentum and effectiveness of some data use teams.

Finally, although teams were trained in the importance of collecting process documentation including the alarm-to-action logs and RTM inventory tools, during the inception training teams were not consistent in completing these. In particular, process logs and meeting minutes were not maintained regularly within any county. Thus, the lack of complete records limited the study team's ability to cross-check observed patterns of refrigerator performance with documented actions at the facility level. Further, in the absence of action log data, the RTM dashboard cannot discern whether an alarm condition ends due to human intervention or other external factors. This limited the ability of the study to measure the effect of county personnel structures and escalation processes on the time taken to resolve cold chain equipment issues and respond to temperature excursions.

CONCLUSION

This study reinforces and expands on previous research by Anderson and Comes, which suggested that a combination of RTM data with improved processes for data review and issue management can lead to important improvements in cold chain performance in resource-

constrained settings.[15,17] The results demonstrated that the real-time alarms for temperature excursions increased staff's awareness of cold chain performance and their responsiveness to temperature excursions. At the same time, the positive trends in equipment uptime indicate that data use teams played a key role in identifying and prioritizing recurring issues and facilitating longer-term solutions. The study suggests that the combination of various stakeholders in the data use teams and the problem solving structures and processes the teams followed enabled issues that may otherwise have gone unnoticed or remained unresolved to be addressed or escalated more effectively. The observed decrease in regularity of data use team meetings after the suspension of funding for those meetings highlights the importance of ensuring continued support for these teams. Although we believe that the observed results strongly indicate the value of this intervention in terms of potential vaccine losses averted, further study in this area is needed.

As the study was not designed to separate the effect of the RTM system from a structured approach for data review and issue management inherent in the data use team processes, further studies are needed to separate out issues that can be effectively solved by the technology alone versus those that require human or behavioral intervention. However, the study does provide evidence to support combining the use of an RTM system with a structured data use and problem solving team approach as a highly beneficial strategy to improve vaccine cold chain performance throughout the supply chain.

Acknowledgments: We acknowledge the government of Kenya, particularly Dr. Collins Tabu, Dr. Joyce Charo, and Catherine Silali; the staff of the National Vaccines and Immunization Program; the county governments; and the health workers in Nairobi, Kajiado, and Isiolo counties for their deep partnership and commitment throughout the work described in this article. We thank all the health workers across Kenya who regularly submitted data to DHIS 2 because these data were instrumental in our work. Our deep gratitude goes to the ELMA Vaccines and Immunization Foundation and the Bill & Melinda Gates Foundation for funding this important work and supporting this research. The authors also acknowledge the reviewers and administrative staff within JSI, inSupply Health, and Nexleaf Analytics who supported the production of this article, as well as the inSupply Kenya team and partners who supported the work described in this article.

Funding: This work was made possible with support from the ELMA Vaccines and Immunization Foundation.

Competing interests: Mr. Ross, Mrs. Kania, and Mrs. Yavari each report grants from JSI during the conduct of the study, other business from UNICEF, and grants from Gavi the Vaccine Alliance and the ELMA Vaccines and Immunization Foundation outside the submitted work.

REFERENCES

1. UNICEF, World Health Organization, World Bank Group, United Nations. *Levels and Trends in Child Mortality Report 2017: Estimates Developed by the UN Inter-agency Group for Child Mortality Estimation*. New York, NY: UNICEF; 2017.
2. Berhane Y, Demissie M. Cold chain status at immunisation centres in Ethiopia. *East Afr Med J*. 2000;77(9):476–479. Medline
3. Andre F, Booy R, Bock H, et al. Vaccination greatly reduces disease, disability, death and inequity worldwide. *Bull World Health Organ*. 2008;86(2):140–146. CrossRef. Medline
4. World Health Organization (WHO). Immunization coverage. WHO website. http://www.who.int/news-room/fact-sheets/detail/immunization-coverage. Accessed August 14, 2018.
5. World Health Organization (WHO). *Temperature Sensitivity of Vaccines*. Geneva, Switzerland: WHO; 2006. http://apps.who.int/iris/bitstream/handle/10665/69387/WHO_IVB_06.10_eng.pdf?sequence=1. Accessed August 14, 2018.
6. World Health Organization (WHO). How To Monitor Temperatures in the Vaccine Supply Chain. Geneva, Switzerland: WHO;2015. http://apps.who.int/iris/bitstream/handle/10665/183583/WHO_IVB_15.04_eng.pdf;jsessionid=CFEA9237632F0C91A1030619C334D81A?sequence=1.
7. Matthias DM, Robertson J, Garrison MM, Newland S, Nelson C. Freezing temperatures in the vaccine cold chain: a systematic literature review. *Vaccine*. 2007;25(20):3980–3986. CrossRef. Medline
8. Techathawat S, Varinsathien P, Rasdjarmrearnsook A, Tharmaphornpilas P. Exposure to heat and freezing in the vaccine cold chain in Thailand. *Vaccine*. 2007;25(7):1328–1333. CrossRef. Medline
9. Nelson C, Froes P, Van Dyck AM, et al. Monitoring temperatures in the vaccine cold chain in Bolivia. *Vaccine*. 2007;25(3):433–437. CrossRef. Medline
10. Nelson CM, Wibisono H, Purwanto H, Mansyur I, Moniaga V, Widjaya A. Hepatitis B vaccine freezing in the Indonesian cold chain: evidence and solutions. *Bull World Health Organ*. 2004;82:99–105. Medline
11. Wirkas T, Toikilik S, Miller N, Morgan C, Clements CJ. A vaccine cold chain freezing study in PNG highlights technology needs for hot climate countries. *Vaccine*. 2007;25(4):691–697. CrossRef. Medline
12. Murhekar MV, Dutta S, Kapoor AN, et al. Frequent exposure to suboptimal temperatures in vaccine cold-chain system in India: results of temperature monitoring in 10 states. *Bull World Health Organ*. 2013;91:906–913. CrossRef. Medline
13. Hanson CM, George AM, Sawadogo A, Schreiber B. Is freezing in the vaccine cold chain an ongoing issue? A literature review. *Vaccine*. 2017;35(17):2127–2133. CrossRef. Medline
14. Yakum MN, Ateudjieu J, Pélagie FR, Walter EA, Watcho P. Factors associated with the exposure of vaccines to adverse temperature conditions: the case of North West region, Cameroon. *BMC Res Notes*. 2015;8(1):277. CrossRef. Medline
15. Anderson R, Perrier T, Pervaiz F, et al. Supporting immunization programs with improved vaccine cold chain information systems. Papers presented at: IEEE Global Humanitarian Technology Conference (GHTC); October 10–13, 2014; San Jose, CA. CrossRef
16. Kenya Ministry of Health. Kenya EVM Assessment: 28 October–30 November 2013. December 2013.
17. Comes T, Bergtora Sandvik K, Van de Walle B. Cold chains, interrupted: The use of technology and information for decisions that keep humanitarian vaccines cool. *J Humanit Logist Supply Chain Manag*. 2018;8(1):49–69.

Peer Reviewed

Received: April 30, 2019; **Accepted:** October 22, 2019; **First Published Online:** December 16, 2019

Cite this article as: Lutukai M, Bunde EA, Hatch B, et al. Using data to keep vaccines cold in Kenya: remote temperature monitoring with data review teams for vaccine management. *Glob Health Sci Pract.* 2019;7(4):585-597. https://doi.org/10.9745/GHSP-D-19-00157

GLOBAL HEALTH: SCIENCE AND PRACTICE
Dedicated to what works in global health programs
OPEN ACCESS

METHODOLOGY

Using a Chord Diagram to Visualize Dynamics in Contraceptive Use: Bringing Data into Practice

Amy Finnegan,[a,b,c] Saumya S. Sao,[b] Megan J. Huchko[b]

A chord diagram is an innovative tool that can be used to visualize switching and quitting in contraceptive use between 2 discrete time points. It complements existing analysis of contraceptive failure rates and provides a richer understanding of contraceptive discontinuation and method switching that can lead to fresh insights to improve family planning programs.

ABSTRACT

We applied a novel approach to visualizing contraceptive histories from the Demographic and Health Surveys (DHS) contraceptive calendar to elucidate patterns of contraceptive switching and discontinuation (e.g., "churn" in contraceptive use across 2 points in time). Taking the contraceptive calendar from the 2014 Kenya DHS, we used R, an open source statistical programming platform, and the *chorddiag* package to create interactive chord diagrams to visualize contraceptive use trajectories, including switching and discontinuation, for all contraceptive types queried in the DHS. We present screenshots from the interactive version. We also tested the usefulness of our chord diagram with potential users, including family planning researchers and advocates. Chord diagrams are visually appealing and provide users with the ability to investigate unique patterns in contraceptive discontinuation using publicly available data from the DHS. After receiving a brief orientation to a chord diagram, users found the chord diagram easy to understand and manipulate. The chord diagram is a potentially powerful way for family planning researchers, advocates, and program managers to visualize women's contraceptive trajectories and provides insights into the churn in contraceptive use between 2 discrete time periods.

INTRODUCTION

In low-income countries, one-third of women who initiate a modern method of contraception discontinue within the first year, and one-half discontinue within the first 2 years, potentially putting them at risk for unintended pregnancies, maternal morbidity, and mortality.[1,2] Dis-continuation rates from hazard models describe the magnitude of the problem, but lack detail on the pathways women take after quitting or switching methods. A better understanding of the reasons for and patterns in contraceptive discontinuation may help identify intervention points to ensure that women who do not desire pregnancy have access to contraceptive methods that meet their family planning needs.

Family Planning 2020 (FP2020) aims to enable an additional 120 million women and girls to use contraceptive methods by 2020. The impact of this ambitious goal will be attenuated if high numbers of women who begin using methods later discontinue. In fact, an estimated 38% of women with an unmet need for family planning are former users of contraception.[3] This phenomenon of discontinuers has been labeled the "leaky bucket"[4]—even when new users begin contraception many of them later quit and total users may decrease. Although some level of contraceptive discontinuation is anticipated based on individual preferences, a better understanding of the rates and reasons for discontinuation among women who do not desire pregnancy will help to more effectively address unmet need.

The Demographic and Health Survey (DHS) collects data on contraceptive discontinuation through the contraceptive calendar module that was first included in DHS surveys conducted between 1988 and 1991, starting with the second wave. The contraceptive calendar is a retrospective monthly reporting of contraceptive use, births, and reasons for discontinuation over the last approximately 5 years from the date of the survey. Although these data provide a more detailed picture of contraceptive behavior, in their raw form, they can be difficult to navigate without advanced data analysis skills. Family planning advocates, program planners, and practitioners could benefit from more explanatory data to develop programs and advocacy campaigns directed to increase the use of contraceptive methods and meet the FP2020[5] and Sustainable Development Goals.[6]

[a] IntraHealth International, Chapel Hill, NC, USA.
[b] Duke Center for Global Reproductive Health at the Duke Global Health Institute, Durham, NC, USA.
[c] Evidence Lab at the Duke Global Health Institute, Durham, NC, USA.
Correspondence to Amy Finnegan (amy.finnegan@gmail.com).

The objective of this article is to describe the use of a chord diagram, a novel data visualization technique, to elucidate contraceptive trajectories among users of contraception as captured through the DHS contraceptive calendar. Ultimately, better visualization can lead to better understanding of "churn" and the "leaky bucket" and contribute to programs that meet women's need for effective family planning.

DATA AND METHODS

Data

DHS surveys provide a cross-sectional snapshot of a country's population of women of reproductive age that is representative at the national, subnational, urban, and rural levels.[7] The Standard DHS is typically collected in 5-year intervals. The DHS has been conducted in 77 countries since the first wave in 1985 until 2017. Beginning in 1990, 65 countries (85%) have included the contraceptive calendar for a total of 168 surveys. Surveys that contain calendars can be identified using the DHS Application Programming Interface and the *rdhs* package in R.[8]

Each Standard DHS survey contains demographic information about women, including their level of education, marital status, fertility preferences, and contraceptive use. For the contraceptive calendar, women are asked to report their contraceptive use and pregnancy status for each month for the 5 years before the survey. Enumerators anchor contraceptive use to events such as the birth of a child or pregnancy. Women who say they have stopped using a contraceptive method are asked to give the reason for discontinuation according to predefined categories. The DHS calendar relies on retrospective reporting, which can lead to underreporting of contraceptive use if use occurred further back in time and may differ by whether a method is user-dependent, such as the pill, or not user-dependent, such as an intrauterine device.[1,9,10]

This article uses the contraceptive calendar from the 2014 Kenya DHS, which included 31,079 women of reproductive age (15–49). Approximately 50% of these women (n=14,741) responded to the "long questionnaire" that included the contraceptive calendar. Although use of another survey could present interesting patterns, we chose the most recent survey from Kenya because the country has made commitments to increase access to contraceptives for women, especially in vulnerable areas, and has a moderate rate of unmet need for contraception (17.5% according to the most recent DHS).[7]

Methods

We created event files from the DHS calendar data where each row represented a person-month. Although we created the event files using R,[11] an open source coding platform, the DHS has created a comprehensive contraceptive calendar tutorial that has code for creating event files in Stata and SPSS.[12] These event files can be read into R and used to create the diagrams we describe in this article. To identify new episodes of contraceptive use after non-use, we created a subset of data to include only contraceptive use that was reported after 1 month of non-use. For example, in January 2014, a woman may have reported using the pill. She was included in the sample if she reported non-use of contraception in December 2013, the month prior. We then created a subset of data to include only the person-months during the first month of use and 12 months later. We used this data to create a matrix that shows transitions in contraceptive use between the first month of reported use (baseline), any reported use after a month of non-use that occurred during the 5-year survey period, and 12 months after an episode of use after 1 month of non-use aggregated across all common trajectories between baseline and 12 months. Only women with reported values at both baseline and 12 months were included in the sample. Women who never reported contraceptive use and those who reported discontinuing a method of contraception because they wanted to become pregnant were excluded so that the focus was on women who quit or switched methods while desiring to avoid becoming pregnant. We included all other reasons for discontinuation in the visualization described in this article, but others could choose to exclude them depending on their research question. We used individual sampling weights provided by the DHS.

Better data visualization can contribute to programs that meet women's needs for effective family planning.

We visualized these trajectories using the *chorddiag* package[13] in R to create an interactive chord diagram that we describe here using static screen shots to show the interactive features. A chord diagram is a circular visualization of interrelated data akin to a transition matrix (shown in the Table), with states represented along arcs and flow between states represented as chords. To visualize contraceptive trajectories from the DHS, we organized the circular chord diagram into halves using the "bipartite" option in the *chorddiag* package. A tutorial of how to read a bipartite chord

A chord diagram is a circular visualization of interrelated data with states represented along arcs and flow between states represented as chords.

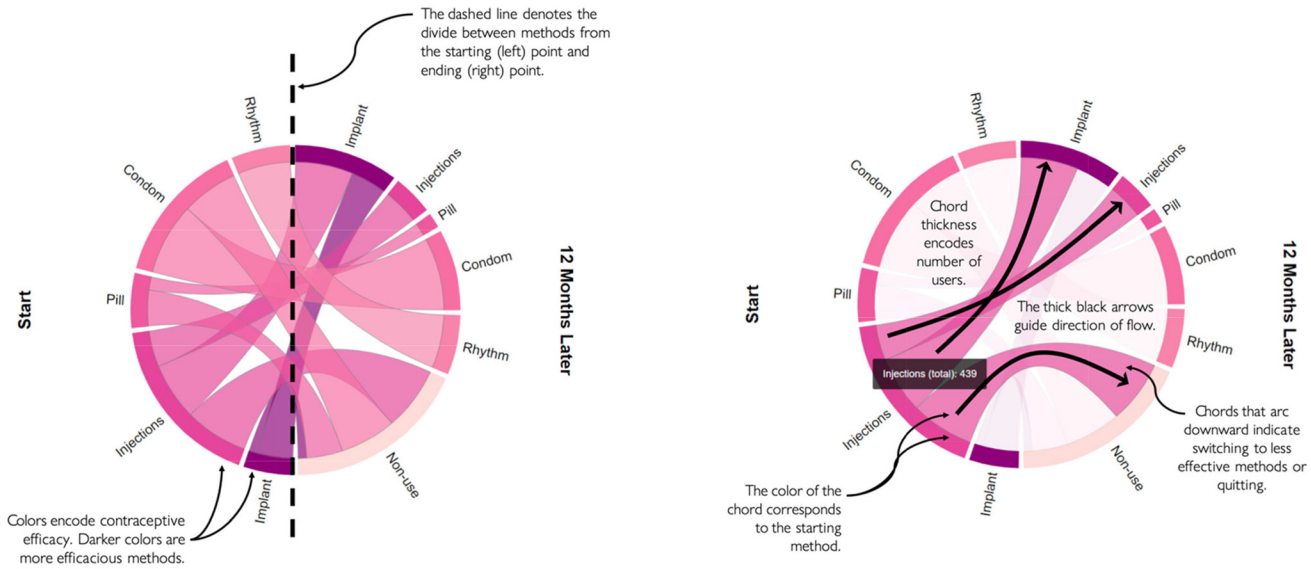
TABLE. Data Matrix of Hypothetical Contraceptive Use Dynamics

Method Used at Baseline	Method Used 12 Months Later					
	Implants	Injections	Pill	Condom	Rhythm	Non-use
Implants	100					24
Injections	161	100				178
Pill			36			100
Condom				202		158
Rhythm					148	

Source: Hypothetical data created for tutorial. See Figure 1 for the chord diagram visualization.

diagram is shown in Figure 1 using the hypothetical population of women shown in the Table. The beginning period is shown on the left half of the circle, and the ending period is shown on the right half of the circle dissected by a dashed line (panel (a) of Figure 1). The starting population and ending population contain the same number of women so the circle is split directly in half by the 2 periods. Chord diagrams encode the size of flows from 1 period to the next by the width of each chord connecting the starting period to the subsequent analysis period. In Figure 1, the chord size represents the number of women who used each method at baseline and follow-up (i.e., 12 months later). The same hypothetical group of women are displayed in the Table. Focusing on injection users (row 2), there were 439 total injection users at baseline. One hundred women who used

FIGURE 1. How to Read a Chord Diagram Showing Contraceptive Use



a. Basic description of the chord diagram

b. Contraceptive trajectory of injection users (interactive feature)

This figure displays a chord diagram using hypothetical data on contraceptive use at 2 time points: baseline (start) and follow-up (12 months later). The dashed line (Panel a) splits the chord diagram between the 2 time points. The interactive feature of the same chord diagram (Panel b) is shown, with a focus on women reporting injection use at baseline. Arrows and dotted lines are added for demonstration purposes and are not present in the actual chord diagrams.

injections at baseline were still using injections 12 months later, 161 switched to implants, and 178 quit using injections and weren't using any method.

The chord diagram (Figure 1) has several features that make it easier to see patterns not immediately obvious in the Table. In the chord diagram, contraceptive methods are organized in order of effectiveness with typical use[14,15]; darker colors indicate higher efficacy (see panel (a) of Figure 1). In the interactive version, the viewer can use the mouse to hover over a contraceptive method to highlight specific flows to or from that method over time and see both the number of women and the directionality. The color of the chord is set to match the starting period so the viewer can easily see the direction of flows. Panel (b) of Figure 1 illustrates that chords that arc downward indicate women who switched to a less effective method or discontinued use of any method, and chords that arc upward indicate women who switched to a more effective method. If women continued their initial method, the chord draws a line the size of the non-switching population to the same method on the right-hand side of the circle.

We chose the *chorddiag*[13] package over other R packages because of its ability to work seamlessly "out of the box" for our purposes and others. Other packages to create Sankey diagrams,[16] another type of flow visualization developed to track energy flows out of systems,[17] did not allow us to keep the order of contraceptive methods by efficacy and instead placed the bars where they created the least amount of overlap. Likewise, *chorddiag* package versions that created net flow visualizations (the "directional" option) that showed each contraceptive method once around the circumference of the chord diagram and displayed net in and outflows from each state on the same arc were ultimately too confusing to understand compared to "bipartite" chord diagrams used in this article.

The idea for a contraceptive trajectory visualization was pilot-tested at a large, family planning research NGO based in Durham, NC, among staff working directly with contraceptive use data, including the contraceptive calendar from the DHS. After receiving feedback on the need for an interactive visualization tool, the research team developed an early prototype of the visualization using data from the 2014 Kenya DHS, the 2014–2015 Rwanda DHS, and the 2014 Guatemala DHS (these countries were chosen because their survey periods occurred the most recently and overlapped) and debuted it at the International Conference on Family Planning (ICFP) 2018 in Kigali, Rwanda. The pilot-testers were able to interact with the web-based tool and provided feedback on whether the chord diagrams were easy to interpret, showed interesting patterns, and would be of use to the family planning community.

RESULTS

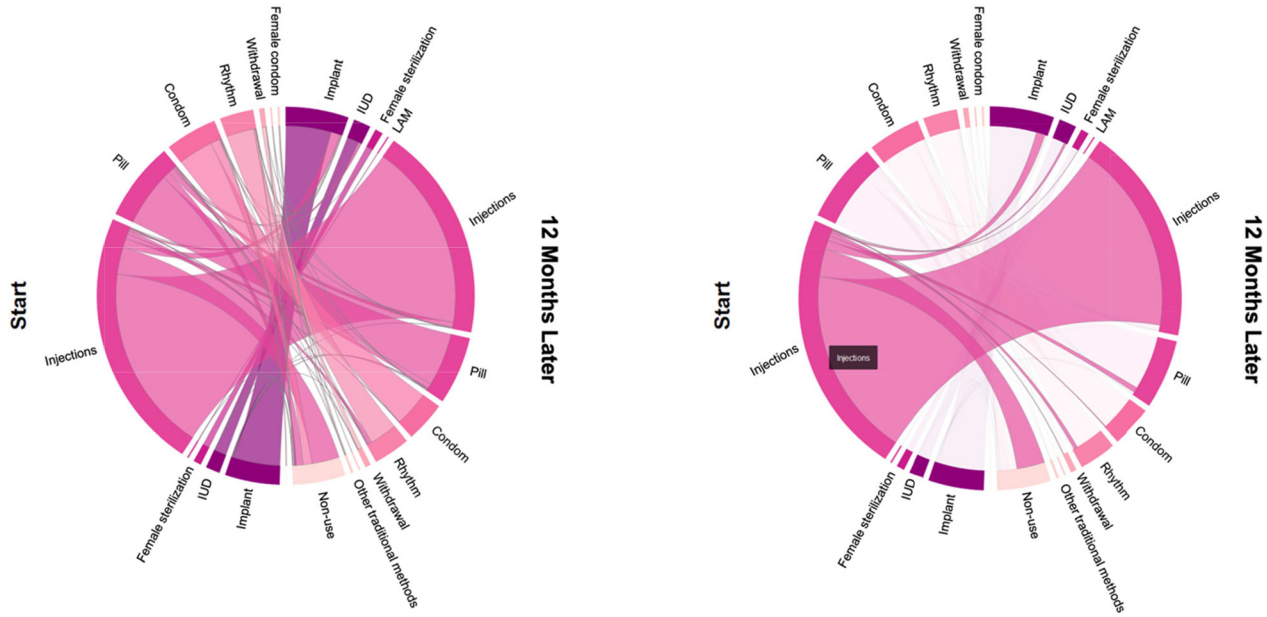
Contraceptive Use Dynamics

We applied chord diagrams to visualize 12-month contraceptive trajectories for women surveyed in the 2014 Kenya DHS contraceptive calendar. The chord diagram in Figure 2 shows the contraceptive use patterns of the 3,783 women in the Kenyan DHS contraceptive calendar who reported new use of a method of contraception (e.g., use of any method of contraception after 1 month of reported non-use). The aggregated contraceptive trajectories between baseline and 12 months are shown in Panel (a) of Figure 2. The size of each chord around the circumference represents the number of users of each method, weighted by the DHS sampling weight. The arc for injections is the largest because they were the most frequently used method, though they were not the most effective method available indicated by their medium dark color. The light-colored region indicating methods with low efficacy grew from baseline to 12 months, showing that transitions from injections tend to be toward less effective methods or, more typically, non-use. Panel (b) of Figure 2 shows a screen capture of a user hovering the mouse over injection users at baseline. Note that although most users were still using injections 12 months later, as indicated by the thick chord connecting one-half of the chord diagram to the other (the same color as the injection arc), more of those who stopped using injections chose methods that were less effective or stopped using contraception (chords that arc downward) than chose methods that were more effective than injections (chords that arc upward).

Chord diagrams can be used to visualize contraceptive quitting and switching to elucidate the reasons for contraceptive use behaviors.

Chord diagrams can also be used to visualize contraceptive quitting and switching to elucidate the reasons for contraceptive use behaviors. We generated a chord diagram to visualize the reasons that women either switched to another method (Figure 3, Panel a) or discontinued use of a method (Figure 3, Panel b). Figure 3 visualizes reasons for discontinuation or switching among the subset of women using any method at baseline but no method or a different method at the 12-month follow-up. The focus of the example interactive feature is again on women who were using

FIGURE 2. Trajectories of New Contraceptive Users Among Women Sampled in the 2014 Kenya Demographic Health Survey



a. Distribution of contraceptive use over 12 months among new contraceptive users

b. Contraceptive trajectory of new injection users over 12 months (interactive feature)

The start period (left) begins with a woman's first reported use after non-use of contraception in the prior month. The right side (12 months later) displays the method she was using, if any, 12 months later. This population of women mostly uses injections between the 2 time periods (Panel a). Most women who quit using injections either switch to less effective methods or stop using contraception altogether. The trajectories of injection users specifically is shown (Panel b). A few women switch to more effective methods, but most stay on injections. Among those who are no longer using injections 12 months later, most have quit using any method of contraception.

injections at baseline. Recall from Figure 2 panel (b) that most women continued using injections at 12 months, and that most women who stopped using injections quit using contraception altogether rather than switch to another method. Panel (a) of Figure 3 shows that among women who switched from injections to some other method, most did so because of side effects. Cost, access, and inconvenience comprised a small portion of reasons for switching. Panel (b) of Figure 3 shows that nearly all of the women who stopped using injections and did not start using another method did so because of side effects. Comparing reasons for switching (a) to reasons for non-use (b) revealed that most women who switched did so for volitional reasons, including wanting more effective methods or methods with fewer side effects. Discontinuation also was not heavily influenced by cost, access, and convenience suggesting that these reasons may not have been as large a contributor to discontinuation as method-specific reasons.

Family planning researchers and practitioners were eager to use these chord diagrams as an innovative visualization of DHS contraceptive calendar data.

User Testing

User testing revealed that family planning researchers and practitioners were eager to utilize these chord diagrams as an innovative visualization of DHS contraceptive calendar data. Users commented on the importance of using visualization tools to allow for development of new and more nuanced inferences about contraceptive "churn" from 1 time period to the next among a population of contraceptive users to meet the FP2020 goals. Users who had previously seen a chord diagram present data on migration without 2 distinct spheres (not showing left vs. right or baseline vs. endline)[18] or other types of flow diagrams (e.g., Sankey diagrams[17]) questioned the value added of the type of chord diagram that showed flows between 2 periods. However, nearly all users were unfamiliar with chord diagrams and were eager to learn more about chord diagrams that could showed flow between 2 periods. These responses demonstrated that the chord diagrams

FIGURE 3. Reasons for Contraceptive Method Discontinuation Among Women Sampled in the 2014 Kenya Demographic Health Survey

a. Reasons for method switching among injection users (interactive feature)

b. Reasons for discontinuation (interactive feature)

These chord diagrams display the reasons for switching to another method (Panel a) or discontinuing (Panel b). Colors along the left-hand side encode method effectiveness. The colors along the right-hand side encode reasons for discontinuation into "in need" (dark blue) and not "in need" (light blue). In both cases, the most common reason women quit using injections was because of side effects (both quitting and switching). Comparing Panel (a) to (b), about half of women who quit using any method were still in need while women who switched were still in need.

may require a brief tutorial or instructive animation to orient users who are new to this visualization method, which is the motivation for this article. After users were provided with such a demonstration, they were able to easily navigate the chord diagram independently.

DISCUSSION

We found that visualization of contraceptive use, switching, and discontinuation as trajectories (e.g., stocks and flows) from 1 period to the next using a chord diagram provides a richer portrait of contraceptive calendar data that better describes women's experiences than calculation of discontinuation or failure rates alone. Users showed keen interest in employing the chord diagram but required a short tutorial on the method before being able to take full advantage of the innovative visualization method.

Chord diagrams are popular for visualizing migration, another demographic quantity characterized by stocks and flows.[18] There are similarities in characterizing migration and contraceptive trajectories, including overall rates of and reasons for migration and differences in those rates and reasons by starting point and destination. Given these similarities, the chord diagram should be a similarly versatile and popular tool to help elucidate trends in contraceptive behaviors.

Although the chord diagram is a powerful method to see overall trends in contraceptive behavior in defined populations and time periods, there are documented caveats to using the contraceptive calendar data to track individual contraceptive use including recall and social desirability bias.[1,9,19] Aggregated statistics appear to suffer from less retrospective recall bias,[20] though this may vary across populations. Since the use of chord diagrams described in this article show aggregated statistics, the data should be no more biased than traditional hazard models. One caveat to consider when viewing a chord diagram of contraceptive use between 2 time points is that the

Visualizing contraceptive use, switching, and discontinuation provides a richer portrait of data that better describes women's experiences than calculating discontinuation or failure rates alone.

number of users who switch or discontinue may be small, therefore, small/thin arcs should not be overinterpreted. One remedy may be to collapse methods into short- and long-acting and look at churn across higher-order categories rather than by efficacy alone.

The chord diagram described here was the best "out of the box" package in R (*chorddiag* package[13]) to create interactive chord diagrams, making it an easy tool to get up and running for users. The contraceptive calendar tutorials[12] from the DHS can be used to create event files; at the time of writing this article programs are only available for Stata and SPSS but more programs are being added, which can be loaded into R and manipulated to create interactive chord diagrams. Once users have created event files for the 168 available contraceptive calendars, they could easily switch between surveys in R and create interactive chord diagrams. Users may be able to think of other demographic or public health quantities than can be visualized as flows with a chord diagram.

CONCLUSION

The chord diagram is a potentially useful way to visualize women's contraceptive trajectories and can complement a single indicator of the rate of contraceptive discontinuation obtained from hazard models. A chord diagram visualization can be used to augment the hazard of discontinuation calculated using DHS data.

This interactive visualization provides a more dynamic look at contraceptive trajectories that, in the hands of practitioners, researchers, and family planning advocates, can help generate new insights into the contraceptive trajectories that women experience throughout their reproductive lives. The ability to visualize a cohort of women's contraceptive decision making in detail has important implications for supply chain, health worker development, budget priorities, and contraceptive guidelines. Better knowledge about country-specific trends and questions will allow family planning programmatic investments to reach more women and girls.

Acknowledgments: The authors wish to acknowledge the helpful contributions from family planning advocates, researchers, and practitioners at the International Conference on Family Planning 2018, researchers at FHI 360, attendees of a brown bag presentation at the Demographic and Health Surveys offices in Rockville, MD, in March 2019, and Kat Tumlinson at the Carolina Population Center in Chapel Hill, NC. This work would not have been possible without financial support from Bass Connections and Data+ at Duke University and all the undergraduate and graduate students who were part of our Big Data for Reproductive Health (bd4rh) research team during academic year 2018–2019.

Funding: Bass Connections and Data+ at Duke University

Competing Interests: None declared.

REFERENCES

1. Bradley SEK, Schwandt HM, Khan S. *Levels, Trends and Reasons for Contraceptive Discontinuation*. DHS Analytical Studies No. 20. Calverton, Maryland: ICF Macro; 2009. https://dhsprogram.com/pubs/pdf/AS20/AS20.pdf. Accessed July 28, 2019.
2. Castle S, Askew I. *Contraceptive discontinuation: Reasons, Challenges, and Solutions*. New York, NY: Population Council and FP2020; 2015. http://www.familyplanning2020.org/resources/contraceptive-discontinuation-reasons-challenges-and-solutions. Accessed July 28, 2019.
3. Jain AK, Obare F, RamaRao S, Askew I. Reducing unmet need by supporting women with met need. *Int Perspect Sex Reprod Health*. 2013;39(3):133–141. CrossRef. Medline
4. Jain AK. The leaking bucket phenomenon in family planning. Champions4Choice website. http://champions4choice.org/2014/09/the-leaking-bucket-phenomenon-in-family-planning/. Published September 9, 2014. Accessed July 28, 2019.
5. Family Planning 2020. https://www.familyplanning2020.org/about-us. Published 2019. Accessed April 11, 2019.
6. United Nations. Sustainable Development Goals Knowledge Platform. Sustainable Development Goals website. https://sustainabledevelopment.un.org/. Published 2019. Accessed April 11, 2019.
7. Kenya National Bureau of Statistics, Ministry of Health, National AIDS Control Council, Kenya Medical Research Institute, National Council for Population and Development. *Kenya Demographic and Health Survey 2014*. Rockville, MD: ICF International; 2015. http://dhsprogram.com/pubs/pdf/FR308/FR308.pdf. Accessed April 11, 2019.
8. Watson OJ, Eaton J. How to use rdhs. Comprehensive R Archive Network website. https://cran.r-project.org/web/packages/rdhs/vignettes/introduction.html. Published January 22, 2019. Accessed July 28, 2019.
9. Bradley SEK, Polis CB, Bankole A, Croft T. Global contraceptive failure rates: who is most at risk? *Stud Fam Plann*. 2019;50(1):3–24. CrossRef. Medline
10. Bradley SEK, Winfrey W, Croft T. *Contraceptive Use and Perinatal Mortality in the DHS: An Assessment of the Quality and Consistency of Calendars and Histories*. DHS Methodological Reports No. 17. Rockville, MD: ICF International; 2015.
11. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2018. https://www.R-project.org. Accessed April 1, 2019.
12. DHS Program. DHS contraceptive calendar tutorial. DHS Program website. https://www.dhsprogram.com/data/Calendar-Tutorial/index.cfm. Accessed July 28, 2019.
13. Flor M. R interface to D3 chord diagrams. Chorddiag. https://github.com/mattflor/chorddiag. Accessed July 28, 2019.
14. Family planning/contraception. World Health Organization website. https://www.who.int/en/news-room/fact-sheets/detail/family-planning-contraception. Published February 8, 2018. Accessed July 28, 2019. [10.1016/j.contraception.2019.07.008]
15. Hatcher RA. *Contraceptive Technology*. 20th ed. New York, NY: Ardent Media; 2011.
16. Bock T. How to create Sankey diagrams from tables (data frames) using R. R-Bloggers website. https://www.r-bloggers.

com/how-to-create-sankey-diagrams-from-tables-data-frames-using-r/. Published August 17, 2018. Accessed July 28, 2019.

17. Schmidt M. The Sankey diagram in energy and material flow management. *J Ind Ecol.* 2008;12(1):82–94. CrossRef
18. Abel GJ, Sander N. Quantifying global international migration flows. *Science.* 2014;343(6178):1520–1522. CrossRef. Medline
19. Callahan RL, Becker S. The reliability of calendar data for reporting contraceptive use: evidence from rural Bangladesh. *Stud Fam Plann.* 2012;43(3):213–222. CrossRef. Medline
20. Strickler JA, Magnani RJ, McCann HG, Brown LF, Rice JC. The reliability of reporting of contraceptive behavior in DHS calendar data: evidence from Morocco. *Stud Fam Plann.* 1997;28(1):44–53. CrossRef. Medline

Peer Reviewed

Received: June 5, 2019; **Accepted:** July 16, 2019; **First Published Online:** December 5, 2019

Cite this article as: Finnegan A, Sao SS, Huchko MJ. Using a chord diagram to visualize dynamics in contraceptive use: bringing data into practice. *Glob Health Sci Pract.* 2019;7(4):598-605. https://doi.org/10.9745/GHSP-D-19-00205

GLOBAL HEALTH: SCIENCE AND PRACTICE
Dedicated to what works in global health programs
OPEN ACCESS

LETTER TO THE EDITOR

Saving Mothers, Giving Life: Don't Neglect the Health Systems Element

Krishna Hort,[a] Louise Simpson[b]

See related articles in the SMGL supplement.

INTRODUCTION

We congratulate the authors of the articles in the GHSP supplement on the Saving Mothers, Giving Life (SMGL) project in Uganda and Zambia. The significant reduction in maternal deaths arising from the project is heartening, and we are pleased to see this comprehensive description of the project, its interventions and outcomes, and a range of studies evaluating its impact, all published in full.

In this letter, we would like to focus on an aspect of the project that we feel did not receive adequate attention in the supplement, namely, its role as a health systems strengthening (HSS) initiative. To do so, we draw on the literature describing the characteristics of HSS initiatives and seek to highlight the HSS elements of the SMGL program based on the articles in the supplement. Given the relatively sparse literature on HSS characteristics, we also draw on our own experience of HSS in relation to maternal health programs through the Australia-Indonesia Partnership for Maternal and Neonatal Health (AIPMNH) in eastern Indonesia, over the period 2009 to 2015 (unpublished).

HEALTH SYSTEMS STRENGTHENING

The SMGL initiative was clearly conceived as an HSS initiative, using a "systems approach, focused at the health district level"[1] and addressing 5 elements of the health system in an integrated manner. This systems approach was designed to "create a highly visible, bold initiative that would galvanize global action and financial support"[2] and demonstrate that such an initiative "could achieve impressive results in a short time."[1]

[a] Nossal Institute for Global Health, University of Melbourne, Melbourne, Australia.
[b] Australia Indonesian Partnership for Maternal and Neonatal Health, Kupang, Indonesia.
Correspondence to Krishna Hort (khort@unimelb.edu.au).

However, literature on HSS emphasizes that it goes beyond simply addressing health system components. HSS involves[3]:

> *investments in inputs in an integrated and systemic way, but also reforming the architecture that determines how different parts of the health system operate and interact to meet priority health needs through people-centered integrated services.*

An HSS approach also takes the complex and adaptive nature of health systems into consideration,[4] which has given rise to the view that HSS is also[5]:

> *a complex, iterative, and learning process wherein the interactions between actors, structures, services, and subsystems are optimized over time while striving for health systems goals.*

Based on this perspective, evaluations of HSS should include the process of implementation to understand how the HSS intervention interacts with and adapts to the operating environment.[4] However, the implementation process is not always well captured in evaluations of HSS interventions, as noted by Adam et al.[4] in their review of studies of HSS.

Unfortunately, the collection of articles on the SMGL initiative has also somewhat neglected the implementation process element. The focus of reporting and evaluation appears to have been on the specific interventions and their links to results and outcomes. As noted by one of the SMGL authors[6]:

> *Although extensive monitoring and evaluation activities were implemented for SMGL, these methods focused heavily on measuring effects on health outcomes and much less on process documentation of various programmatic approaches.*

The minimal focus on process documentation occurred despite the fact that[7]:

> *the majority of the interventions supported by SMGL were not "new" to the host country; rather, they were existing interventions that were refined, strengthened, and, in most cases, taken to greater scale of implementation through partnership.*

Consequently, the implementation process seems to have in fact been an important element of the SMGL initiative. With this in mind, we reviewed the articles to try

to identify where and how the project adapted to different contexts, responded to contextual changes, and evolved during the implementation process.

LEVELS OF IMPLEMENTATION

To identify the implementation elements, we use the framework proposed by Samuels et al.[8] and the approach taken by Cleary et al.[9] to describe an HSS project in Mozambique. Samuels et al.[8] referred to 3 levels of implementation—macro-level governance, meso-level partnerships, and micro-level local ownership.

Macro Level

At a macro level, Samuels et al.[8] identified the following factors as being key to implementation: effective governance, coherent evidence-based policies, and partnerships between donors and national-level actors that encourage the latter to take control and enable transition from donor funds to national funding streams. Cleary et al.[9] also identified the importance of "relational trust building" with partners in terms of interpersonal and institutional trust as contributing to implementation success.

The SMGL articles refer to implementation at a macro level in managing the multiple partnerships involved and in gaining national government support and commitment. They report some success in building high levels of ownership among district health leaders, as well as "increased [Ministry of Health] commitment and heightened social awareness" at the national level with regard to the prevention of maternal and newborn deaths.[2] However, the papers contain little information on how these relationships were built.

One challenge to these relationships appears to have been the funding "gap" that occurred between funding for the initial year and a decision on whether to continue funding for a longer period.[2] Erratic funding in subsequent years was also mentioned. Thus, project implementers needed to manage the "frustration" among government and implementing partners generated by these funding requirements. Learning how this management was accomplished would be interesting.

SMGL used a funding mechanism designed to encourage increasing national government contribution through progressive annual reductions in funds provided.[2] However, the impacts of this process on relationships were not discussed, although the failure to gain increased government funding mentioned by Healey et al.[7] suggests that such impacts may well have contributed a further implementation challenge.

Our experience in Indonesia also demonstrates the critical importance of developing partnership with key government managers; building partner government capacity in leadership and governance, which requires long-term commitment (5 years plus in our case); and ensuring close collaboration in decision making.

Meso Level

Samuels et al.[8] described the meso level as the level at which:

> *policies become specific interventions shaped by organizational structures and procedures, and partnerships among different organizations.*

Such interventions occur primarily at a subnational level and are closely related to the capacity and context of district administration, for example, the extent of decentralization.

SMGL implementation at a district level relied on multiple partners, including teams of Ugandan and Zambian government medical and local civic leaders as well as "equally dedicated and talented U.S. government teams."[1] This approach was reported to have enabled[1]:

> *considerable problem solving, resource gathering, and resilience in the face of unexpected administrative and logistical challenges.*

In addition, the approach was reported to have contributed to addressing both supply- and demand-side barriers that accelerated change.[2]

The SMGL articles describe some implementation challenges associated with this approach. For example, in Zambia, the tools and systems for facility data collection were developed separately by individual partners and were not harmonized across districts; consequently, some indicators could not be aggregated at baseline.[10] In addition, the management of multiple partnerships created a heavy administrative burden, and the SMGL authors recommended a smaller partnership in the future.[2]

The meso level is also where implementation might encounter contextual changes and need to adapt to changing circumstances. The SMGL authors referred to changes in the district structure in both countries and to adaptations during implementation to address issues of ambulance sustainability, transport voucher demand, and scope, such as the expansion to include postpartum and neonatal care.[2] However, the articles

contain no reference to the extent or implications of decentralization, which operates in both countries and represents an important contextual factor.

The articles address the need to develop capacity at the district level in "planning, execution, and evaluation,"[2] and to the use of national intermediaries to support implementation. SMGL district coordinators—often retired midwives—were hired to harmonize all SMGL activities in their district with district health officers and district health management teams, and to serve as a link with implementing partners.[2]

Our experience in the AIPMNH was similar. Although district governments in Indonesia operated in a highly decentralized environment, district capacity in planning and execution was low. We found that the provision of flexible funding from the project enabled district governments to introduce new approaches and activities that could not be funded through the complex government planning and budget process. However, planners needed training and support in the planning and execution process to ensure proposed activities were appropriate and effectively implemented.

Micro Level

At a micro level, the SMGL articles describe extensive use of community extension workers to engage with communities and to provide a bridge between health services and pregnant women and their families. The workers included village health team volunteers in Uganda and Safe Motherhood Action Groups in Zambia. In both countries, the workers advocated for birth preparedness, promoted healthy practices, and encouraged antenatal care visits, facility deliveries, and postpartum care.[6]

We also used this strategy in Indonesia, building on and strengthening existing village and community institutions, such as the health post (posyandu) concept to support safe pregnancy and delivery, and engaging with church groups, which wield considerable influence in this largely Christian area.

One of the key challenges for these community-level activities is ensuring sustainability, and we were heartened to see that Healey et al.[7] commented that the formalization and institutionalization of the community volunteer groups was one of the most significant signs of sustainability.

Learning and Adaptation

The SMGL papers also imply a learning and adaptive process during implementation, with Conlon et al. referring to the development of a "think tank" atmosphere toward the latter stages of the project.[2] This process is of considerable interest to us, because we also noted how the AIPMNH program evolved from a focus on implementation toward support for the development and conduct of studies, interventions and evaluations of innovative practices, and the exchange of this information across districts and nationally.

DISCUSSION

The implementation experience of SMGL, although not well documented, supports many of the aspects identified by Cleary et al.[9] regarding the HSS project in Mozambique. They noted the need to develop ownership, build trust, adapt to contextual change, and have long-term adaptive support.

However, it is worth noting that Cleary et al.[9] stressed the need for long-term funding commitment and for flexibility in scheduling, for example, through lengthy start-up phases.[9] This area appears to be one in which the SMGL approach of initial "conditional" funding for the start-up year, followed by a 6-month "chaotic" period waiting for a decision on implementation, was not in line with best practice in HSS and resulted in a recommendation that[2]:

> *any future systems approach should commit to a minimum of 5 years support from the outset.*

Finally, we note that, although choosing the right strategies and interventions is important in a systems approach, the implementation process and managing implementation can be just as important. In a journal devoted to science and practice, let us not neglect the "practice" element.

This point was well made by Cleary et al.[9]:

> *However, implementation practice in HSS is rarely reported in close detail, and these "obvious" issues are rarely intentionally managed, reported, or measured. This case study shows how important implementation practice can be as it underpins HSS intervention activities and their success—and suggests that it may need to be taken more seriously into account by funders, intervention designers, implementers, and researchers—as a key element of intervention design, management, and evaluation.*

Acknowledgments: The Australia-Indonesia Partnership for Maternal and Neonatal Health was managed by Coffey International Ltd. and funded by the Australian Department of Foreign Affairs and Trade. The managers and funders had no role in the comments made in this article.

Competing Interests: None declared.

REFERENCES

1. Quam L, Achrekar A, Clay R. Saving Mothers, Giving Life: a systems approach to reducing maternal and perinatal deaths in Uganda and Zambia. *Glob Health Sci Pract.* 2019;7(suppl 1):S1–S5. CrossRef. Medline
2. Conlon CM, Serbanescu F, Marum L, et al. Saving Mothers, Giving Life: it takes a system to save a mother. *Glob Health Sci Pract.* 2019; 7(suppl 1):S6–S26. CrossRef. Medline
3. UHC2030. *Healthy Systems for Universal Health Coverage—A Joint Vision for Healthy Lives*. Geneva: World Health Organization and the World Bank; 2017 https://openknowledge.worldbank.org/handle/10986/29231. Accessed July 21, 2019.
4. Adam T, Hsu J, de Savigny D, Lavis JN, Røttingen JA, Bennett S. Evaluating health systems strengthening interventions in low-income and middle-income countries: are we asking the right questions? *Health Policy Plan.* 2012;27:iv9–iv19. CrossRef. Medline
5. Rockefeller Foundation; Health Systems Action Network; African Centre for Global Health and Social Transformation; Doris Duke Charitable Foundation. Strengthening health system capacities through institutional development: enhancing collaboration between donors and organizations in low-income countries; August 27–31, 2012; Bellagio, Italy. Final meeting report. http://ghsia.files.wordpress.com/2013/06/bellagio-meeting-final-report.pdf. Accessed July 21, 2019.
6. Serbanescu F, Goodwin MM, Binzen S, Morof D, Asiimwe AR et al. Addressing the first delay in Saving Mothers, Giving Life districts in Uganda and Zambia: approaches and results for increasing demand for facility delivery services. *Glob Health Sci Pract.* 2019;7(suppl 1): S48–S67. CrossRef. Medline
7. Healey J, Conlon CM, Malam K, Hobson R, Kaharuza F et al. Sustainability and scale of the Saving Mothers, Giving Life approach in Uganda and Zambia. *Glob Health Sci Pract.* 2019;7(suppl 1): S188–S206. CrossRef. Medline
8. Samuels F, Amaya AB, Balabanova D. Drivers of health system strengthening: learning from implementation of maternal and child health programmes in Mozambique, Nepal and Rwanda. *Health Policy Plan.* 2017;32(7):1015–1031. CrossRef. Medline
9. Cleary S, Erasmus E, Gilson L, et al. The everyday practice of supporting health system development: learning from how an externally led intervention was implemented in Mozambique. *Health Policy Plan.* 2018;33(7):801–810. CrossRef. Medline
10. Morof D, Serbanescu F, Goodwin MM, et al. Addressing the third delay in Saving Mothers, Giving Life districts in Uganda and Zambia: ensuring adequate and appropriate facility-based maternal and perinatal health care. *Glob Health Sci Pract.* 2019;7(suppl 1):S85–S103. CrossRef. Medline

Received: May 16, 2019; **Accepted:** June 30, 2019; **First Published Online:** October 17, 2019

Cite this article as: Hort K, Simpson L. Saving Mothers, Giving Life: don't neglect the health systems element. *Glob Health Sci Pract.* 2019;7(4):606-609. https://doi.org/10.9745/GHSP-D-19-00178

GLOBAL HEALTH: SCIENCE AND PRACTICE
Dedicated to what works in global health programs
OPEN ACCESS

LETTER TO THE EDITOR

Authors' Response to "Saving Mothers, Giving Life: Don't Neglect the Health Systems Element"

Florina Serbanescu,[a] Claudia Morrissey Conlon,[b] Frank Kaharuza,[c] Masuka Musumali[d]

See related articles by Hort and Simpson and in the SMGL supplement.

On behalf of the Saving Mothers, Giving Life (SMGL) Technical Working Group, we would like to thank Holt and Simpson[1] for their valuable insights. We also appreciate their interest in finding out more about the planning and implementation process of the initiative, the challenges encountered, and the adaptations needed to account for contextual factors in each country.

We agree with the authors on the importance of bringing attention to the valuable experiences and lessons learned during the implementation and course of the SMGL activities, including the activities directly related to health systems strengthening (HSS). We generally agree with the points raised regarding a lesser focus on the evaluation of the process of implementation of the SMGL initiative in the supplement in favor of highlighting the outcomes and impacts of the initiative on maternal and newborn health. However, we beg to differ that the implementation experiences of SGML are "not well documented."

The articles constituting the supplement do not represent an exhaustive account of all aspects of the SMGL initiative. The select articles published in the supplement have largely focused on the outcomes of the initiative at its conclusion after 5 years of implementation. They add to already published accounts about the initial planning, implementation, and monitoring and evaluation of the SMGL interventions, including a comprehensive external evaluation of inputs and processes undertaken during the first year of the initiative.[2] They also add to the article by Kruk and colleagues that focused on the effects of the SMGL on the health systems in Uganda and Zambia during Phase 1.[3]

In the context of describing extraordinary, effective, multisectoral, and large-scale interventions that reduced maternal mortality in the SMGL-supported districts, the supplement includes numerous examples of HSS. The articles describing the comprehensive district system strengthening approaches that led to reductions in the "Three Delays" give ample details about strategies employed at the individual, community, health facility, and district levels. Successes and challenges to implementation of these strategies and increased accountability demanded by the initiative are thoroughly documented.[4] The SMGL model builds on an integrated approach with complex converging factors that have contributed to its success. These include a well-functioning public-private partnership, country leadership, integration into and strengthening of the existing health systems, resource mobilization, community participation, and commitments to rigorous monitoring and evaluation. The outcomes and impacts presented in the supplement were agreed-upon tracer indicators selected before the launch of the initiative and designed to capture the main effects on the maternal and child health status. The full contributions of SMGL to the health systems strengthening and the wellbeing of communities in Uganda and Zambia are not entirely amenable to quantitative monitoring and evaluation.

The fact that the authors were able to identify examples in the supplement to discuss SMGL's successes at the macro, meso, and micro levels attests to the wealth of implementation details provided by the articles in the supplement. We echo the value of examining HSS through a more structured and formalized lens, though "there is little consensus on what health systems strengthening (HSS) entails, what the drivers of successful HSS initiatives are, and how they can be measured."[5] We concur that a set of well-defined and agreed-upon guiding principles and indicators, similar to those of SMGL, are very important for monitoring and evaluation of any complex health initiative or HSS.

We thank Holt and Simpson for noting the importance of HSS in global health programming and research and recognize the value of continuing to share the experiences and lessons learned during the course of planning,

[a] Division of Reproductive Health, Centers for Disease Control and Prevention, Atlanta, GA.
[b] United States Agency for International Development (USAID), Saving Mothers, Giving Life, Washington, DC.
[c] USAID, Kampala, Uganda.
[d] USAID, Lusaka, Zambia.
Correspondence to Florina Serbanescu (fxs7@cdc.gov).

implementing, and evaluating the SMGL initiative. We echo their thoughts on the need for future research to tease out more firmly the critical components of HSS in Uganda and Zambia. There is a wealth of qualitative and quantitative evidence that captured these crucial experiences. Continued analyses and documentation of these aspects may include bringing forward country- and district-level insights and experiences of the SMGL initiative related to HSS. Uganda and Zambia have already embarked on a road of scaling up components of the SMGL model. Policy makers and program managers in other low- and middle-income settings where similar approaches could be used to rapidly reduce maternal mortality may greatly benefit from learning about the SMGL's role in improving health systems.

Competing interests: None declared.

REFERENCES

1. Holt K, Simpson L. Saving Mothers, Giving Life: Don't Neglect the Health Systems Element. *Global Health Sci Pract.* 2019; 7(3): CrossRef
2. Saving Mothers, Giving Life website. http://savingmothersgivinglife.org. Accessed September 5, 2019.
3. Kruk ME, Rabkin M, Grepin KA, et al. 'Big push' to reduce maternal mortality in Uganda and Zambia enhanced health systems but lacked a sustainability plan. *Health Aff (Millwood).* 2014;33(6):1058–1066. CrossRef. Medline
4. Conlon CM, Serbanescu F, Marum L, et al. Saving Mothers, Giving Life: it takes a system to save a mother. *Glob Health Sci Pract.* 2019;7 (suppl 1):S6–S26. CrossRef. Medline
5. Samuels F, Amaya AB, Balabanova D. Drivers of health system strengthening: learning from implementation of maternal and child health programmes in Mozambique, Nepal and Rwanda. *Health Policy Plan.* 2017;32(7):1015–1031. CrossRef. Medline

Received: September 10, 2019; **First Published Online:** October 17, 2019

Cite this article as: Serbanescu F, Conlon CM, Kaharuza F, Musumali M. Authors' response to "Saving Mothers, Giving Life: don't neglect the health systems element." *Glob Health Sci Pract.* 2019;7(4):610-611. https://doi.org/10.9745/GHSP-D-19-00313

Global Health: Science and Practice

Knowledge SUCCESS
Johns Hopkins Bloomberg School of Public Health
Center for Communication Programs
111 Market Place, Suite 310
Baltimore, MD 21202
Phone: 410-659-6134
Fax: 410-659-6266
Email: editorialoffice@ghspjournal.org
Web: www.ghspjournal.org