RESEARCH ARTICLE

Advancing contraceptive security, availability, and choice in Malawi using a quality improvement methodology [version 1; peer review: awaiting peer review]

Jewel Gausman¹, Megan E. Schmitt ², Jane Wickstrom³

¹Harvard T H Chan School of Public Health, Cambridge, MA, USA
²EngenderHealth, Inc, New York, NY, USA
³Bill & Melinda Gates Foundation, Seattle, WA, USA

Abstract

Many initiatives to improve contraceptive security (CS) rightly focus on strengthening national and regional systems. However, local health facilities are often under-resourced and lack technical capacity that feed into the larger supply chain. This study’s objective was to assess whether changes in facility CS indicators were associated with participation in 2014-2016 implementation and scale-up of a quality improvement methodology—Client-Oriented, Provider efficient services (COPE®) for Contraceptive Security—in 60 facilities across 10 districts of Malawi. The intervention included facility self-assessment guides and action plans to address local challenges.

Results showed significant improvements in facilities having both a trained provider and contraceptive supplies. The percentage of health centers with all requirements for implant services increased significantly, including implant removal (from 26.5%; 95% CI: 14.9-41.1 to 77.6%; 95% CI: 63.4-88.2, p<.001). Health centers (from 0.0%; 95% CI: 0.0-7.3 to 10.2%; 95% CI: 3.4-22.2, p<0.05) and hospitals (from 45.5%; 95% CI: 16.7-76.7 to 90.9%; 95% CI: 58.7-99.8, p<0.05) significantly improved in the percentage of facilities able to insert intrauterine devices. Hospitals improved their ability to offer female sterilizations (27.2%; 95% CI: 6.0-61.0 to 63.6%; 95% CI: 30.7-89.1, p<0.05) and male sterilizations. Low performing health centers showed significant improvement in staff capacity, logistics management information systems, equipment, and total CS performance. The percentage of facilities placing emergency orders for contraceptives during the three months prior to an assessment showed a decreasing, non-significant trend among hospitals but was significant among health centers (from 69.2%; 95% CI: 54.6-81.7 to 36.7%; 95% CI: 23.4-51.7; p<0.001). Facility staff commitment was associated with action item completion. Improvements tended to be sustained over time. Community engagement is thought to be important to intervention success. COPE for CS may be an effective intervention and future research/programs can build off of this preliminary programmatic experience when seeking to address last mile challenges.
Keywords
contraceptive security, family planning, reproductive health, quality improvement, health systems, logistics and supply chain management, service delivery, Malawi, Sub-Saharan Africa

Corresponding author: Megan E. Schmitt (megschmittmsph@gmail.com)

Author roles: Gausman J: Conceptualization, Data Curation, Formal Analysis, Methodology, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; Schmitt ME: Conceptualization, Data Curation, Methodology, Project Administration, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; Wickstrom J: Conceptualization, Funding Acquisition, Project Administration, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: Ms. Schmitt and Ms. Wickstrom report grants to EngenderHealth, Inc. from the U.S. Agency for International Development, PATH on behalf of the Reproductive Health Supplies Coalition, and the Bill & Melinda Gates Foundation to support the conduct of the study. Ms. Wickstrom was an employee of EngenderHealth, Inc. during the design, implementation, data collection, analysis and drafting of the manuscript. She is currently an employee of the Bill & Melinda Gates Foundation, which partially funded the analysis of study data.

Grant information: The RESPOND project was made possible by the generous support of the American People through the U.S. Agency for International Development (USAID), under the terms of the cooperative agreement GPO-A-000-08-00007-00. Support for the Malawi Contraceptive Security Project was provided by PATH, on behalf of the Reproductive Health Supplies Coalition (NOD.1841-705853-GRT). Additional support for data analysis was provided by the Bill & Melinda Gates Foundation (OPP1084384). The views expressed here do not necessarily reflect those of USAID or the United States Government, the RHSC or PATH, nor the Bill & Melinda Gates Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2019 Gausman J et al. This is an open access article distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Gausman J, Schmitt ME and Wickstrom J. Advancing contraceptive security, availability, and choice in Malawi using a quality improvement methodology [version 1; peer review: awaiting peer review] Gates Open Research 2019, 3:1111 (https://doi.org/10.12688/gatesopenres.12896.1)

First published: 03 Apr 2019, 3:1111 (https://doi.org/10.12688/gatesopenres.12896.1)
Introduction

Contraceptive security (CS) exists when people are able to choose, obtain, and use the contraceptive methods and services they desire from among a full range of methods (see Box 1)\(^1\). Achieving CS is critical to meeting the Sustainable Development Goals, especially Goal 3, which seeks to ensure healthy lives and promote well-being for all at all ages\(^2\). The international public health community recognizes that CS remains weak in many resource poor settings of sub-Saharan Africa and elsewhere, calling for action and commitments in two recent London Summits on Family Planning in 2012 and 2017\(^3\)-\(^5\). Importantly, improving availability and choice of contraceptive methods and services is essential to fulfilling sexual and reproductive health and rights.

Improving CS requires systems transformation and concerted efforts to make improvements sustainable. Many donor and government-funded initiatives aimed at improving CS rightly focus on strengthening national and perhaps regional-level systems of forecasting, procurement, central stock management, supply chain and related elements. However, district zones and their local health facilities that are closer to clients are often under-resourced and/or lack technical capacity in logistics management, requisition, stock management and stock reporting which feed into national systems. These “last mile” facilities face challenges that require specific tools and approaches designed to identify and solve local problems as part of the larger supply chain\(^6\). Examples of important initiatives geared towards last mile needs include: tools for community-based distribution programs; community score cards for accountability and community involvement in CS; and, initiatives in pharmacy management\(^7\)-\(^9\). Recognizing gaps in efforts for the last mile, the Reproductive Health Supplies Coalition’s (RHSC) Advocacy and Accountability Work Group announced a call to join its new “Last Mile Workstream” as recently as late 2017\(^10\).

Recognizing a gap in methodologies, tools and approaches specifically targeted to facility management of CS issues at the last mile, EngenderHealth’s RESPOND project developed and tested the quality improvement methodology COPE\(^®\) for Contraceptive Security (COPE stands for Client-oriented, Provider Efficient services)\(^11\). The methodology includes facility self-assessment guides and subsequent development and implementation of facility-level action plans to address local gaps and challenges to contraceptive security. This article presents findings from the 2014–2016 implementation and scale-up of the methodology in Malawi with support of the RESPOND project and the RHSC. The initiative was carried out at the request of, and in partnership with, the Ministry of Health and district health officials in 60 facilities across 10 districts of Malawi.

Program description

COPE\(^®\) for Contraceptive Security

Client-oriented, provider-efficient services (COPE\(^®\)) is a quality improvement methodology first developed in 1995 by EngenderHealth to address clients’ rights to health services and provider needs to deliver quality services\(^12\)-\(^13\). Since that time, numerous iterations of COPE for different technical areas have been tested and published, including adaptations to improve the quality of services in reproductive health, HIV care and treatment, male circumcision, and abortion care, among others\(^14\)-\(^18\). The COPE\(^®\) for Contraceptive Security methodology and tools are used by frontline health and logistics personnel to identify and implement low-cost, local solutions to address problems related to contraceptive supply\(^19\)-\(^20\). The process incorporates staff accountability and linkages with district supervision systems and community and local leadership, as needed, creating ownership in improving quality and strengthening systems for sustainability.

The COPE for CS process begins with an exercise conducted by trained facilitators to orient facility teams on the activity. Once staff agree to tackle the issues under consideration, facility teams complete a series of self-assessment guides on issues ranging from stock management, reporting, requisition, transportation, warehousing and personnel. Problem identifications lead to staff developing action plans to address local bottlenecks and issues that the facility and district can try to address themselves, and formation of a COPE for CS committee to oversee implementation and follow-up of their action plan. A job aid is available to foster continued reflection and reanalysis of issues during implementation and for use in district supervision\(^21\). Intended for global use, COPE for CS was originally designed and tested in Tanzania from 2011 to 2013, where results showed statistically significant improvements in contraceptive availability and increases in family planning use after more than one year\(^1,22\).

Introduction and scale-up in Malawi

In Malawi, the Ministry of Health (MOH) made remarkable progress in improving family planning access over the past 15 years. Modern contraceptive prevalence among married women increased from 28% in 2004 to 42% in 2010, and again to 58.1% by 2015–2016\(^23\)-\(^25\). The MOH made strides toward achieving contraceptive security at the national level, while noting that contraceptive security is weaker at district and lower-level health facilities. In 2014, the MOH Directorate for Reproductive Health (DRH) requested assistance to address contraceptive security at the last mile. Challenges identified at the local level included: lack of trained providers (especially for long-acting reversible contraception (LARCs) and permanent methods); unclear roles and responsibilities for logistics management; and, lack of training in requisitioning and ordering (which can lead to

---

**Box 1.**

Contraceptive security exists when people are able to choose, obtain, and use the contraceptive methods and services they desire from among a full range of methods (short-acting, long-acting reversible, and permanent). In order for family planning programs to provide a full range of methods, three basic elements must be consistently present at a service delivery point: the contraceptives themselves; necessary medical equipment, instruments, and expendable supplies; and trained staff able to provide each method. When any of these elements is missing from a service delivery point, the method cannot be offered, and contraceptive security is neither achieved nor maintained\(^1\).
Implementation process of COPE for CS

To understand the intervention, it is necessary to understand the structure of the COPE for CS process. Site facilitators, trained in the methodology and tool, lead a facility team through an initial exercise and their continuous quality improvement efforts. COPE for CS is designed to be easy to implement without outside technical assistance and is adaptable for local facility contexts. The site facilitators, in coordination with facility leadership, determine the length of their initial COPE for CS facility exercise based on workflow at the facility, usually consisting of several hours in the late afternoon when client flow is slower over two to three days. In larger facilities, staff from multiple departments are asked to join the exercise, including those outside family planning and/or logistics management. For example, facilities are encouraged to invite cleaning personnel who use supplies for infection prevention and guards who may be a client’s first point of contact at the facility. In smaller facilities, the entire staff may work on COPE for CS together for the duration of the work. Whatever staff configuration is chosen, the goal is to have teams with first-hand operational experience with different types of challenges within the facility and who want to participate in problem solving with their colleagues, generating a shared sense of ownership for results.

Self-assessment guides

The COPE for CS tool includes 10 self-assessment guides containing a series of questions, based on international standards, regarding the quality of services, systems, and procedures (see Box 3). Site facilitators and teams review the guides during the initial exercise and can complete individual assessments as a team, in small groups, pairs, or as individuals, depending on their preference and the staff participating. Each guide includes instructions on which type of staff, by function, is best placed to respond to its questions. The guides are designed to be flexible and adaptable to the facility team’s needs. Staff can write in issues that are not directly raised by a guide and are relevant, and can choose to skip questions they do not find relevant to their context. In addition, the team is not required to complete all 10 assessments at one time; rather, they can prioritize which guides to use and at which points in their exercises/process. After completing guide(s), the facility team reviews and identifies issues at the site as revealed by assessment questions.

Action plans

Following discussion of any one of the self-assessment guides, facility teams develop an action plan to consolidate and prioritize recommendations. Action plans identify problems related to CS; identify the root cause(s) of each problem; propose action items that are realistic, measurable, attainable, and address the root cause(s); assign an individual facility staff member responsibility for each action item; ensure a time-bound goal for stock-outs of contraceptives and related supplies). Box 2 shows the programmatic process for COPE for CS introduction and implementation in the two districts as well as 2015–2016 implementation in additional districts and facilities at the request of MOH/DRH.

Box 2. Introduction & scale-up of COPE for CS in Malawi

2014

The MOH Directorate for Reproductive Health requested technical assistance to address contraceptive security at the last mile.

- January 2014: MOH/DRH convened a partner coordination meeting with EngenderHealth RESPOND staff, local implementing partners and donors to introduce the COPE for CS methodology, tools and plans for implementing the methodology.
- January–February 2014: Tanzanian master trainers who co-designed, launched, and evaluated the original COPE for CS trained site facilitators from 18 public-sector health facilities in Mangochi and Salima districts, promoting South-to-South learning.
- February 2014: Site facilitators led facility teams through COPE for CS exercises at each of the 18 facilities. Local project staff served as resources, as needed.
- August 2014: EngenderHealth project staff conducted follow up visits 18 preliminary sites. Qualitative data collection via key informant interviews noted improvements in stock management, on-time ordering, decreases in stock-outs, and improvements in collaboration between facilities and district medical stores.

2015–2016

The Malawi COPE for Contraceptive Security Project continued supporting the preliminary 18 sites and introduced the intervention at 42 scale-up sites (totaling 60 sites across 10 out of 28 districts nationwide).

- February 2015: With the MOH/DRH convened a national-level meeting and orientations for national and district health authorities, donors, and implementing partners to review gain consensus on the initiative.
- April–May 2015: COPE for CS scale up launched with two training-of-trainer (TOT) events for 27 master trainers and 13 district and national supervisors (10 District Family Planning Coordinators and three supervisors from MOH/DRH).
- May–June 2015: Participants in the TOTs then trained site facilitators from the 60 supported sites (three trainings with 20 participants each).
- June–September 2015: Site facilitators led COPE for CS facility exercises at their respective sites, initially supported by master trainers to ensure quality and fidelity to the intervention. For scale-up sites this was the first introduction of COPE for CS to a facility team, while preliminary facilities conducted refresher exercises and revised their action plans originally developed in 2014.
- August–October 2015: EngenderHealth project staff conducted follow up visits at both preliminary and scale-up sites to report on progress, troubleshoot, and document for broader learning.
- June 2015–January 2016: District Health Management Team staff agreed to incorporate discussion of COPE for CS action plan progress into their regular supervisory visits.
completion of each action item; and provide space to comment on action item status and result. Box 4 shows an excerpt from a study facility’s action plan.

COPE for CS emphasizes targeting problems with root causes at the facility level and within facility control. Facilities are also encouraged to include action items that address problems at the district level, if teams can identify a pathway through which the facility may affect change. For example, a facility requiring more trained staff in logistics management, or more trained providers in FP provision, may advocate to district-level management to assign additional personnel to the facility in need. National problems identified are not recommended for inclusion as facility action items.

COPE for CS committees
In addition to completing self-assessment guides and developing action plans, facility teams form a COPE for CS Committee to ensure follow-up and monitor action plan progress. Opportunities to discuss action plan progress include: regular staff meetings, special committee meetings, and during district supervision. COPE for CS Committees may decide to conduct additional full team COPE for CS exercises and continue to complete self-assessments as new issues arise. The COPE for CS job aid is another resource for staff to revisit key self-assessment issues in an abbreviated manner.

Committee members are encouraged to post the facility’s action plan in a visible place where all staff and the public can see it to show the site is dedicated to quality improvement, to encourage accountability and transparency, and to monitor progress against goals. Committees and facility leadership are encouraged to share their action plans with local stakeholders, health advisory committees, politicians, implementing partners and community organizations, as an advocacy tool to request assistance and resources.

Intervention follow-up and support
The introduction and scale-up projects supported training of trainers, site facilitator trainings, facility exercises, and limited follow up visits, to check on action plan progress and provide space for COPE for CS Committees or facility leadership to ask questions about the methodology or seek implementation guidance as needed.

However, the intervention did not include additional inputs to improve contraceptive security at the 60 sites. The COPE for CS initiative purposefully did not provide technical assistance for clinical or logistics training, for example, nor for the procurement of contraceptives or required equipment and supplies. The idea is for facility staff to look for local solutions. The 10 supported districts received varying levels of support for health services, including FP, from other multilateral agencies and partners. Following COPE for CS exercises and action plan development, the facilities may clearly articulate their needs to district leadership who coordinate donor funding in the decentralized Malawian healthcare system.

Outcomes of interest/research questions
The overall objective of this study is to assess whether improvements in facility performance on contraceptive security indicators is associated with participation in the COPE for CS intervention by comparing baseline and endline performance levels. Primary sub-objectives of this study include to: 1) describe implementation characteristics of intervention components; 2) assess whether facilities achieve key intervention-related outputs; and 3) examine changes in performance according to intermediate and ultimate outcomes, in particular, how implementation of the intervention relates to changes in performance, and whether any improvements in performance observed are sustainable over time. Figure 1 presents a logical framework that illustrates how the programmatic components fit together with the research questions of interest.

Data collection
Design and sampling
We obtained data for this analysis through facility surveys designed to collect facility-level data on performance related to contraceptive security.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Cause(s)</th>
<th>Recommendation</th>
<th>By Whom</th>
<th>By When</th>
<th>Completed?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Untimely submission of reports and requisition of commodities</td>
<td>Lack of guidelines on submission of reports and requisition of commodities and who is the responsible officer</td>
<td>Develop or identify guidelines on submission of reports requisition of commodities, including responsible officer, and paste them on wall</td>
<td>[Individual’s name redacted]</td>
<td>10 September 2015</td>
<td>Yes, on 1 October 2015</td>
</tr>
</tbody>
</table>
In 2014 the Malawian MOH/DRH purposively selected 18 facilities from 2 districts (Salima and Magochi) to participate in the study. In 2015, the MOH/DRH selected 42 additional sites in 8 additional districts (Balaka, Chikwawa, Lilongwe, Mzimba North and South, Nkhotakota, Ntchew, and Thyolo) for the scale-up phase, increasing the total number of study facilities to 60. The MOH/DRH considered districts for inclusion if they reported stockouts in a high percentage of facilities and had previously submitted requests for assistance with contraceptive security. Within these districts, the MOH/DRH and district
medical officers selected which facilities would introduce the COPE for CS methodology using the same criteria.

Data collection occurred in three-waves as described in Figure 2. For all facilities, baseline data collection occurred before the start of the intervention (in 2014 for preliminary facilities and 2015 for scale-up facilities). Data was also collected at preliminary facilities in 2015 as a midline assessment. Endline data collection occurred for all sites in 2016. In the 18 preliminary facilities the COPE for CS intervention was introduced between the 2014 and 2015 assessments and continuously implemented through 2016. In the scale-up facilities, the COPE for CS intervention was introduced between the 2015 and 2016 assessment points.

Survey administration and data management
We used a standardized, facility questionnaire to assess each of the dimensions of contraceptive security, as defined by COPE® for contraceptive security: An assessment guide (RESPOND, 2013). These dimensions included:

1. Organization and staffing;
2. Logistics, management, and information systems (LMIS);
3. Procurement/requisition/stockouts;
4. Inventory control procedures, warehouse and storage;
5. Medical equipment, instruments, and expendable supplies.

EngenderHealth staff trained the data collection team on survey administration. For each of the survey rounds, a one-day data collector training was held during which data collectors reviewed the questionnaire and staff provided instructions on its administration. This training also instructed and assessed understanding of standard precautions for protecting human subjects, and included role-playing exercises.

Data collectors conducted the facility survey in English using a paper questionnaire. Interviewers identified the Facility In-Charge or their designate in each of the facilities to obtain permission to conduct the assessments. Following informed consent procedures, they administered the facility questionnaire. The facility survey tools are included as extended data to this paper.

One data collector implemented the facility assessment at each facility. At the conclusion of each assessment, data collectors sent the completed questionnaire to EngenderHealth project staff, who reviewed the questionnaire to ensure it was complete.

A project staff member trained in appropriate coding and data entry techniques entered data from the physical questionnaires into a Microsoft Excel (Microsoft Corporation, Redmond, Washington, USA) database. A project manager then reviewed data entry, comparing the physical questionnaires to the database and documenting any discrepancies and subsequent discussion and resolution.

Protection of human subjects
The 2014, 2015 and 2016 survey protocols received ethical approval from an EngenderHealth review board. External review was not obtained given that the research did not meet the threshold to qualify as research conducted among human subjects. The data were primarily collected to inform program decision making and questions were not focused on individual perspectives; rather, they focused on health facility capacity, staffing, stock and related issues. Data collectors conducted facility assessments only after administering a standard informed consent form. We employed standard measures to maintain confidentiality and anonymity for the facility staff respondent. All information collected was strictly confidential and used only for study purposes. Respondent names were not stored with the final clean data.

Description of variables
We identified variables included in the analysis according to the components of the logic model presented in Figure 1.

Inputs
Input variables include those related to how the intervention was implemented, such as the number of days spent on exercises, frequency of group discussion of the COPE for CS action plans, and the number of COPE for CS committee meetings in each facility.

Outputs
We developed variables pertaining to action plan quality, content, implementation, and commitment. Project staff assessed action plan quality according to a scoring rubric developed a priori (see extended data) according to several quality dimen-

![Figure 2. Overall facility sampling procedures.](image-url)
sions including: whether the problems were clearly identified, whether the action plan identified root causes, whether the action plans offered attainable/realistic solutions, whether individuals were assigned responsibility, and whether items that were identified were time bound. Staff conducted action plan content mapping to determine whether a facility identified items relating to staffing, LMIS, procurement/requisition, inventory control procedures/receiving supplies, warehousing and storage, transport and distribution, financing/budgeting, planning, and medical equipment/instruments/expendable supplies. Box 5 further details the development of variables related to action plan quality.

We measured commitment to COPE for CS action plans by the frequency of group discussion of the action plan and the number of COPE for CS committee meetings. Project staff assessed completion of items in a facility’s action plan based on facility reporting. After reviewing reported plan updates, project staff calculated the number of action plan items that were not initiated, initiated but not yet complete, and completed at endline.

Outcomes

Outcome variables are divided between intermediate and ultimate outcomes. Intermediate outcomes relate to changes in facility performance according staffing, LMIS, supplies/equipment, storage, and procurement, and were developed based on existing literature and expert consultation. We developed a detailed composite score in order to assess changes in facility performance (Box 6). Ultimate outcomes pertain to performance measures that may be indicative of improvement in overall facility performance. These indicators were adapted from the RHSC’s harmonized list of CS indicators, including number of emergency orders for contraceptives that a facility placed in the three months prior to assessment and the number of contraceptive methods available at a facility on the day of the survey.

We defined method availability as whether the facility had the commodity in stock, all required method-specific equipment, and a provider who is trained in provision (and removal, if applicable) of a given method.

Confounding variables

Due to the possibility of confounding variables to influence both the implementation of the intervention and the outcomes of interest, we stratified results according to a variety of important facility-level variables on which data were available including facility location (urban/rural), facility type (health center or hospital), region, and baseline facility performance.

Data analysis

We analyzed the quantitative data using Stata® v14.0, and produced graphics using Statistical Software R’s® ggplot2 package. We presented descriptive statistics for the aforementioned variables of interest, and stratified according to possible confounding variables. We report means and medians for continuous variables and proportions for dichotomous and categorical variables.

We stratified results by key confounding variables, including baseline performance. We considered facilities that had baseline performance scores over 90% as having limited room for improvement between baseline and endline. We constructed a dichotomous baseline performance variable to stratify facilities that performed below/above the 90% threshold. We used t-tests to assess differences between groups for continuous variables.
chi-2 tests to assess differences in binary variables, and ANOVA to assess differences in continuous outcomes between categorical variables with three or more categories.

To assess changes in proportions between baseline and endline, we used McNemar’s chi-2 paired tests of proportions. We also used simple linear (for continuous outcomes) and logistic (for binary outcomes) regression analysis to assess changes between baseline and endline according to key variables of interests.

As discussed above, a facility’s baseline measurement is the first assessment (2014 for the 18 preliminary facilities and 2015 for the 42 scale-up facilities) and a facility’s endline assessment is the last assessment conducted (2016 for all facilities). For the 18 preliminary facilities that had a midline assessment conducted in 2015, we conducted an additional sub-analysis to examine CS performance trends over time at all assessment points.

Results

Table 1 presents a description of the facility characteristics and details on the types of intervention-related characteristics, by facility type. Hospitals have a greater median number of staff (n=376) as compared to health centers (n=27). Health centers are almost universally located in rural areas (97.6%), as compared to hospitals (45.5%, rural areas).

As shown in Table 2, health centers and hospitals also implement the COPE for CS similarly with regard to the number of days used for the initial exercise (3 days) and the number of CS Committee Meetings per facility (5 meetings). Similarly, most facilities had monthly group discussions of the COPE for CS action plan (health centers, 65.3%; hospitals, 63.6%).

The majority of facilities implemented the initial exercise in three days (75.6% of facilities), while 6.7% used less than three days and 18.7% used more than three days (data not shown). All facilities established a COPE for CS committee and all posted their action plan in a visible space within the facilities (data not shown).

Baseline facility performance according to the five CS dimensions appeared similar across facility type, location, and region (Table 3). Mean total performance score across all facilities was 81.8 (out of 100) at baseline, with facilities performing the highest on staffing (92.9) and lowest on equipment (59.8). Health facilities had significantly lower overall performance scores at baseline than hospitals (80.7 versus 87.0, p<0.01), and specifically with LMIS (87.3 versus 90.9, p<0.05), procurement (80.1 versus 93.1, p<0.01) and equipment (56.8 versus 73.3, p<.001), respectively.

Table 4 presents characteristics of the facility action plans, by facility type. Staffing/organizational capacity is the most common priority dimension identified in action plans (98.3%), while equipment and procurement tend to be the least commonly identified priorities (65.0% and 33.3%, respectively). Health centers were significantly more likely to identify equipment as a priority in their action plans than hospitals (71.4% versus 36.4%, p<0.05).

Results indicate that action plans tended to be of a high quality, with little variation across facilities (the overall mean quality score was 19/20). The mean percentage of items completed per facility was significantly lower in health centers (46.7%, CI 0.0–91.1) versus hospitals (66.8%, CI 12.5–100.0, p<0.05). On average, hospitals completed 20.1% more items in their action plan items than health centers (95% CI: 4.0%, 36.3%; p=0.016) (data not shown). Additionally, health centers also initiated fewer items than hospitals. Simple linear regression analysis revealed no significant differences according to location (rural/urban) and region in the overall quality score of the action plans or the progress made in completing action items.

Results suggest that the frequency of group discussion of the COPE for CS action plan is associated with the facility’s priority in their action plans than hospitals (71.4% versus 36.4%, p<0.05).

Table 1. Description of health facilities at baseline (2014 and 2015).

<table>
<thead>
<tr>
<th>Facility Characteristics</th>
<th>Facility Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health Center (n=49)</td>
<td>Hospital (n=11)</td>
</tr>
<tr>
<td>Med, IQR</td>
<td>Med, IQR</td>
</tr>
<tr>
<td>Number of Staff</td>
<td>27 (22, 44)</td>
</tr>
<tr>
<td>Location</td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>97.6%</td>
</tr>
<tr>
<td>Region</td>
<td></td>
</tr>
<tr>
<td>Central</td>
<td>42.9%</td>
</tr>
<tr>
<td>Northern</td>
<td>18.3%</td>
</tr>
<tr>
<td>Southern</td>
<td>38.8%</td>
</tr>
</tbody>
</table>

Table 2. Description of intervention implementation characteristics among all facilities.

<table>
<thead>
<tr>
<th>Intervention Characteristics</th>
<th>Med, IQR</th>
<th>Med, IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time the facility spent to implement initial COPE for CS orientation in days (IQR)</td>
<td>3 (3, 3)</td>
<td>3 (3, 3)</td>
</tr>
<tr>
<td>Frequency of group discussion of COPE for CS Action Plan</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>More than once per month</td>
<td>14.3</td>
<td>18.2</td>
</tr>
<tr>
<td>Monthly</td>
<td>65.3</td>
<td>63.6</td>
</tr>
<tr>
<td>Less than monthly</td>
<td>18.4</td>
<td>18.2</td>
</tr>
<tr>
<td>No group discussion</td>
<td>2.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table 3. Characteristic of the facility action plans, by facility type.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Median number of COPE for CS Committee Meetings per facility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health Center</td>
<td>5 (4, 6)</td>
</tr>
<tr>
<td>Hospital</td>
<td>5 (4, 5)</td>
</tr>
</tbody>
</table>

Med: Median

IQR: Inter-quartile range
Table 3. Comparison of differences in baseline CS performance score total and by CS dimension, by facility type, location and region.

<table>
<thead>
<tr>
<th>Facility Characteristics</th>
<th>Total Performance Score</th>
<th>Performance Score by CS Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=60</td>
<td>Staff (Mean (SD) p-value)</td>
</tr>
<tr>
<td></td>
<td>n=60</td>
<td>LMIS (Mean (SD) p-value)</td>
</tr>
<tr>
<td></td>
<td>n=60</td>
<td>Storage (Mean (SD) p-value)</td>
</tr>
<tr>
<td></td>
<td>n=60</td>
<td>Procurement (Mean (SD) p-value)</td>
</tr>
<tr>
<td></td>
<td>n=60</td>
<td>Equipment (Mean (SD) p-value)</td>
</tr>
<tr>
<td><strong>Number of Facilities</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health Center</td>
<td>80.7 (6.9) **</td>
<td>92.7 (17.8) **</td>
</tr>
<tr>
<td>Hospital</td>
<td>87.0 (6.7) **</td>
<td>92.8 (7.9) **</td>
</tr>
<tr>
<td><strong>Location</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>81.2 (7.2)</td>
<td>92.6 (17.1) **</td>
</tr>
<tr>
<td>Urban</td>
<td>86.6 (5.5)</td>
<td>95.2 (6.6) **</td>
</tr>
<tr>
<td><strong>Region</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Central</td>
<td>81.5 (9.1)</td>
<td>87.6 (22.9) **</td>
</tr>
<tr>
<td>Northern</td>
<td>80.8 (3.6)</td>
<td>97.5 (5.6) **</td>
</tr>
<tr>
<td>Southern</td>
<td>82.6 (6.3)</td>
<td>96.5 (7.4) **</td>
</tr>
<tr>
<td>**All Facilities Combined</td>
<td>81.8 (7.2) **</td>
<td>92.9 (16.2) **</td>
</tr>
</tbody>
</table>

Notes:
* p-value<0.05; ** p-value<0.01; *** p-value<0.001.

1 Differences according to facility characteristic evaluated with t-tests.

2 Differences according to facility characteristic evaluated with ANOVA.

3 Scores calculated out of a total of 100 points.
### Table 4. Differences in action plan content, quality, and progress by facility type.

<table>
<thead>
<tr>
<th>Action Plan Characteristics</th>
<th>Facility Type</th>
<th>P-value</th>
<th>Overall Total (n=60)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Health Center (n=49)</td>
<td>Hospital (n=11)</td>
<td></td>
</tr>
<tr>
<td><strong>Action Plan Content</strong></td>
<td>Med (range)</td>
<td>Med (range)</td>
<td>Med (range)</td>
</tr>
<tr>
<td>Number items identified per facility</td>
<td>9 (4, 23)</td>
<td>7 (2, 12)</td>
<td>8 (2, 23)</td>
</tr>
<tr>
<td>Percentage of facilities that identified the following areas as priorities in Action Plans</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Staffing/Organizational capacity</td>
<td>97.1</td>
<td>100.0</td>
<td>98.3</td>
</tr>
<tr>
<td>LMIS</td>
<td>79.6</td>
<td>81.2</td>
<td>80.0</td>
</tr>
<tr>
<td>Procurement</td>
<td>30.6</td>
<td>45.5</td>
<td>33.3</td>
</tr>
<tr>
<td>Warehousing</td>
<td>77.6</td>
<td>63.7</td>
<td>75.0</td>
</tr>
<tr>
<td>Equipment</td>
<td>71.4</td>
<td>36.4</td>
<td>* 65.0</td>
</tr>
<tr>
<td><strong>Action Plan Quality</strong></td>
<td>Med (range)</td>
<td>Med (range)</td>
<td>Med (range)</td>
</tr>
<tr>
<td>Overall Quality Score</td>
<td>19 (12, 20)</td>
<td>19 (18, 20)</td>
<td>19 (12, 20)</td>
</tr>
<tr>
<td>Quality of Action Plans by Dimension</td>
<td>(4 points possible per dimension)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problems Clearly Identified</td>
<td>4 (3, 4)</td>
<td>4 (3, 4)</td>
<td>4 (3, 4)</td>
</tr>
<tr>
<td>Root Causes</td>
<td>4 (3, 4)</td>
<td>4 (3, 4)</td>
<td>4 (4, 4)</td>
</tr>
<tr>
<td>Attainable/Realistic</td>
<td>4 (2, 4)</td>
<td>4 (3, 4)</td>
<td>4 (2, 4)</td>
</tr>
<tr>
<td>Individuals Assigned Responsibility</td>
<td>4 (0, 4)</td>
<td>3 (3, 4)</td>
<td>4 (0, 4)</td>
</tr>
<tr>
<td>Time Bound</td>
<td>4 (0, 4)</td>
<td>4 (3, 4)</td>
<td>4 (0, 4)</td>
</tr>
<tr>
<td><strong>Action Plan Progress</strong></td>
<td>% (range)</td>
<td>% (range)</td>
<td>% (range)</td>
</tr>
<tr>
<td>Mean percentage of items completed per facility</td>
<td>46.7 (0.0, 91.1)</td>
<td>66.8% (12.5, 100.0)</td>
<td>* 50.4 (0.0, 100.0)</td>
</tr>
<tr>
<td>Mean percentage of items initiated or completed per facility</td>
<td>56.9 (0.0, 100.0)</td>
<td>73.9 (12.5, 100.0)</td>
<td>* 60.0 (0.0, 100.0)</td>
</tr>
<tr>
<td>Mean number of items completed per facility</td>
<td>4.8 (0, 11)</td>
<td>6.1 (1, 15)</td>
<td>5.0 (0, 15)</td>
</tr>
<tr>
<td>Mean number of items initiated or completed per facility</td>
<td>5.7 (0, 11)</td>
<td>6.5 (1, 15)</td>
<td>5.9 (0, 15)</td>
</tr>
</tbody>
</table>

Notes:
- *p-value<0.05
- Differences according to facility characteristic evaluated with Wilcoxon Rank Sum test
- Differences according to facility characteristic evaluated using t-tests

Completion of action plan items. Simple linear regression analysis (not shown) found that facilities that reported having group discussion of their action plans more than once per month completed on average 21.5% (95% CI: 4.0%, 39.0%; p=0.017) more items than those that discussed the plans either monthly or less than once per month. No significant associations were found between the number of official COPE for CS committee meetings reported or the length of time spent on the initial exercise and the percentage of items completed and/or started.

Table 5 shows whether priority identification in action plans was consistent with areas of low performance. Overall, facilities did not consistently identify areas of low performance as priorities in their action plans. Only 60.6% of low performing dimensions were identified as priorities. Health centers were significantly more likely than hospitals to identify low performing dimensions, 63.2% (95% CI: 57.8–68.7) versus 49.1% (95% CI: 32.7–65.4, p<0.05). Similarly, rural facilities identified a greater percentage of low performing dimensions as priorities when compared to urban facilities (p<0.05). Only one in three (29.2%) health centers with low performance in procurement and only 36.6% of hospitals with low performance in equipment identified the respective dimensions as priorities. We found no significant associations between the percent of low performing
Table 5. Differences in baseline performance levels and identification in action plans according to facility characteristic.

<table>
<thead>
<tr>
<th>Facility Characteristics</th>
<th>Percentage of all dimensions that received a low score</th>
<th>Percentage of low-scoring dimensions identified as priority</th>
<th>Percentage of Low Scoring Dimensions Identified as Priorities in Action Plans</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Facility Type</td>
<td>Number of Facilities</td>
<td>Health Center</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>% (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>Staffing</td>
<td></td>
<td>14</td>
<td>100.0</td>
</tr>
<tr>
<td>Percentage of Low Scoring Dimensions Identified as Priorities in Action Plans</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LMIS</td>
<td></td>
<td>32</td>
<td>68.0</td>
</tr>
<tr>
<td>Storage</td>
<td></td>
<td>30</td>
<td>29.2</td>
</tr>
<tr>
<td>Procurement</td>
<td></td>
<td>46</td>
<td>60.0</td>
</tr>
<tr>
<td>Equipment</td>
<td></td>
<td>60</td>
<td>71.4</td>
</tr>
</tbody>
</table>

Notes:
P: p-value
*p: p-value < 0.05

Differences according to facility characteristics evaluated with chi-squared tests

dimensions identified as priorities and the length of time used for the initial COPE for CS exercise.

Figure 3 compares mean baseline and endline performance scores by CS dimension. Overall, the change in scores suggests improvement between baseline and endline in each individual CS dimension as well as total performance score. This holds true when results are disaggregated by facility type. The improvements between baseline and endline are most pronounced when results are presented separately for low performing facilities. The improvements in low performing health centers reach statistical significance in several areas, including staff capacity, LMIS, equipment, and in total score. We observe qualitative improvement to indicate positive change among hospitals, although the results do reach statistical significance.

Table 6 compares baseline and endline performance with regard to FP commodity stock levels, method-specific trained
providers, required equipment/supplies for specific methods, and placement of emergency orders. Health centers generally show some improvement in the types of contraceptive commodities in stock on the day of the assessment (except for a slight decrease in percentage of health centers that have progestin only pills and female condoms), although none of the differences are large enough to reach statistical significance. Hospitals perform very well at both baseline and endline, and show an increase in the percentage with CycleBeads in stock, though the increase is not statistically significant at the 0.05-level (p=0.08). The percentage of health centers with at least five modern methods in stock on the day of the assessment increased slightly (from 87.8%; 95% CI: 75.2–95.3 to 91.8%; 95% CI: 80.3–97.7) though not statistically significant. All hospitals at both baseline and endline were found to have at least five modern methods in stock. The vast majority of health centers (>90%) had providers trained in general family planning and implant insertion/removal, while all hospitals had providers trained in these skills. The percentage of health centers with a provider trained in interuterine device (IUD) insertion/removal increased

| Table 6. Baseline/endline comparison of FP commodity stock, method-specific trained providers, equipment/supplies, and emergency orders for contraceptives. |
|----------------------------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| **Facilities with the following methods in stock on the day of assessment** |
| **Health Center** | **Baseline** | **Endline** | **Difference** | **Hospital** | **Baseline** | **Endline** | **Difference** |
| Male Condoms | 85.7 (72.5–94.1) | 89.8 (77.8–96.6) | 4.1 (-10.6, 18.7) | 100.0 (71.5–100) | 100.0 (71.5–100) | 0.0 (-9.1, 9.1) |
| Female Condoms | 75.5 (61.1–86.7) | 67.3 (52.4–80.0) | -8.2 (-27.0, 10.7) | 100.0 (71.5–100) | 90.1 (58.7–99.8) | -9.1 (-35.1, 17.0) |
| Fertility Awareness Beads/Cycle Beads | 53.1 (38.2–67.5) | 55.1 (40.2–69.3) | 2.0 (-15.5, 19.6) | 72.7 (39.0–93.9) | 100.0 (71.5–100) | 27.2 (-0.08, 0.62) |
| Progestin Only Pills | 79.5 (65.6–89.8) | 73.5 (58.9–85.1) | -6.1 (-25.6, 13.3) | 100.0 (71.5–100) | 90.1 (58.7–99.8) | -9.1 (-35.1, 17.0) |
| Combined Oral Contraceptive Pills | 87.9 (75.2–95.3) | 93.9 (83.1–98.7) | 6.1 (-4.7, 17.0) | 100.0 (71.5–100) | 100.0 (71.5–100) | 0.0 (-9.1, 9.1) |
| DMPA injectable | 89.8 (77.8–96.7) | 95.9 (86.1–99.5) | 4.1 (-9.2, 17.4) | 100.0 (71.5–100) | 100.0 (71.5–100) | 0.0 (-9.1, 9.1) |
| **Range of methods in stock on day of assessment** |
| % of Facilities that have at least 5 modern methods available | 87.8 (75.2–95.3) | 91.8 (80.3–97.7) | 4.1 (-9.2, 17.4) | 100.0 (71.5–100) | 100.0 (71.5–100) | 0.0 (-9.1, 9.1) |
| **Facility has a provider trained in...** |
| General family planning | 97.8 (89.1–99.9) | 93.8 (83.1–98.7) | -4.1 (-14.0, 5.9) | 100.0 (71.5–100) | 100.0 (71.5–100) | 0.0 (-9.1, 9.1) |
| Implant |
| Insertion | 91.8 (80.3–97.7) | 95.9 (86.1–99.5) | 0.22 (0.4, 0.29) | 100.0 (71.5–100) | 100.0 (71.5–100) | 0.0 (-9.1, 9.1) |
| Removal | 91.8 (80.3–97.7) | 95.9 (86.0–99.5) | 4.1 (-0.1, 0.1) | 100.0 (71.5–100) | 100.0 (71.5–100) | 0.0 (-9.1, 9.1) |
| IUD |
| Insertion | 40.4 (26.3–55.7) | 69.3 (54.5–81.7) | 29.7** (9.1, 50.5) | 90.9 (58.7–99.8) | 81.9 (41.2–97.8) | -9.1 (-48.5, 30.4) |
| Removal | 39.1 (25.1–54.6) | 69.3 (54.5–81.7) | 28.2** (7.3, 49.1) | 90.9 (58.7–99.8) | 81.9 (41.2–97.8) | -9.1 (-48.5, 30.4) |
| Male Sterilization*** |
| Female Sterilization*** |

---

**Notes:**
- *p* < 0.05
- **p** < 0.01
- ***p*** < 0.001
significantly between baseline and endline by nearly 30 percentage points (p<0.05), while there was a slight decrease (not significant) in IUD providers at hospitals. At endline, just over 80% of hospitals had a trained provider in IUD insertion and removal. There was no change observed in the percentage of hospitals with a provider trained in male or female sterilization (45.0%; 95% CI: 16.7–76.6, and 81.8%; 95% CI: 48.2–97.7, respectively).

When examining whether a facility had both a trained provider and all of the essential equipment and supplies (including the contraceptive) to provide a method, the results indicate significant improvements between baseline and endline. The percentage of health centers with everything needed to insert and remove an implant increased significantly between baseline and endline, especially the percentage of health centers able to offer implant removal (from 26.5%; 95% CI: 14.9–41.1 to 77.6%; 95% CI: 63.4–88.2, p<0.001). Both health centers (from 0.0% (95% CI: 0.0–7.3 to 20.4% (10.2–34.3) and hospitals (from 45.5% ; 95% CI: 16.7–76.7 to 63.6%; 95% CI: 30.8–89.1, p<0.05) showed significant improvement in the percentage of facilities able to insert an IUD. Similar, positive trends were found for removals. Hospitals also improved in their ability to offer female sterilizations (27.2%; 95% CI: 6.0–61.0 to 63.6%; 95% CI: 30.7–89.1, p<0.05) and male sterilizations, although the latter was not statistically significant. Finally, the overall percentage of facilities placing emergency orders during the three months prior to the assessment showed a decreasing, but non-significant trend among hospitals but was significant among health centers (from 69.2%; 95% CI: 54.6–81.7 to 36.7%; 95% CI: 23.4–51.7; p<0.001).

Table 7 examines in more detail the extent to which emergency orders were placed in the three months prior to the assessment. The results suggest an overall reduction in the placement of emergency orders across all variables related to facility characteristics (facility type, region, and location), though results were not significant with the exception of the Southern region found to be statistically significant (from 1.3; 95% CI: 0.41–1.85 to 0.4; 95% CI: (-0.0–0.8 on average; p<0.05). In terms of intervention characteristics, the mean number of emergency orders decreased overall (from 1.2; 95% CI: 0.7–1.8 to 0.6; 95% CI: 0.2–1.0, p=0.09), but was only significant among facilities that had group discussion of their action plans more than once
Table 7. Mean emergency orders placed in the three months prior to assessments, by facility and intervention characteristics (n=37).

<table>
<thead>
<tr>
<th>Facility Characteristics</th>
<th>Mean Number of Emergency Orders Placed in 3 Months Prior to Assessment (95% CI)</th>
<th>Health Facility</th>
<th>Baseline</th>
<th>Endline</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Facility Type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health Center</td>
<td>1.3 (0.62–1.95)</td>
<td>0.7 (0.2–1.15)</td>
<td>0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospital</td>
<td>1.0 (-0.32–2.3)</td>
<td>0.3 (-0.52–1.19)</td>
<td>0.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Location</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>1.1 (0.6–1.5)</td>
<td>0.6 (0.2–1.1)</td>
<td>0.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>2.4 (-2.4–7.1)</td>
<td>0.4 (-0.7–1.5)</td>
<td>0.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Region</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Central</td>
<td>1.6 (0.2–2.9)</td>
<td>0.9 (-0.0–1.9)</td>
<td>0.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northern</td>
<td>0.9 (-0.1–1.8)</td>
<td>0.5 (-0.4–1.4)</td>
<td>0.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Southern</td>
<td>1.3 (0.41–1.85)</td>
<td>0.4 (-0.0–0.8)</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Intervention Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency of group discussion of COPE for CS Action Plan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group discussion less than once per month</td>
<td>1.2 (0.6–2.0)</td>
<td>0.8 (0.3–1.3)</td>
<td>0.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group discussion more than once per month</td>
<td>1.3 (0.4–1.8)</td>
<td>0.0 (0.0–0.0)</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>1.2 (0.7–1.8)</td>
<td>0.6 (0.2–1.0)</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:

37 facilities had complete data at baseline and endline on emergency orders. This includes 31 health centers and 6 hospitals.

per month (from 1.3 to 0.0 on average; p=0.01). Overall, among the 37 facilities with complete data on emergency orders at both baseline and endline, there was a mean reduction of 0.8 orders over the last three months (p=0.09); while not statistically significant, this shows a meaningful decrease programatically. The reason for the smaller sample size with regard to this question is that 18 facility respondents at baseline and 4 at endline reported not knowing the number of emergency orders placed in their facilities; these facilities are not included in this calculation to avoid missing data bias.

Figure 4 shows the results of a sub-analysis to assess changes in performance scores by CS dimension between 2014 and 2016 for the 18 preliminary facilities, and between 2015 and 2016 for the scale-up facilities. The chart stratifies by baseline performance level (lower performing facilities and all facilities combined). Generally, the results show that performance continues to improve across most dimensions nearly two years after initial implementation of the intervention, except in the areas of stock and storage. We observed the most pronounced improvements in equipment in the initial facilities between 2015 and 2016.

Discussion

The results highlight four important findings related to the association between both the implementation of the COPE for CS intervention and how the intervention may relate to changes in facility performance across the dimensions of contraceptive security.

Facilities overall maintained fidelity to the intervention

The data available on implementation provide important insight on which components worked well and which could improve in the future. The majority of facilities implemented the
intervention according to plan, meaning they held exercises, developed action plans, displayed action plans in a visible location in the facilities, assembled CS committees, had broader group discussions among their staff. Despite the leeway given to facilities to implement in the way that they thought to be most helpful, we observed little variation in implementation. For example, nearly all facilities held the initial exercise in three days. Additionally, the facility action plans tended to be of high quality based on the *a priori* rubric developed to evaluate them. Given the limited variation observed in quality, however, it is difficult to assess the relationship between the quality of a facility’s action plan and changes in CS performance. Future research may build on the quality rubric presented here to be more discerning. One issue to note is that the quality rubric does not consider whether a facility distinguished areas of low baseline performance and specifically identified action plan items related to those areas.

The data also point to areas for future implementation research and possible improvement. Facilities identified few action plan items in the domains of procurement and equipment, despite low performance in these areas. It is possible that facilities determined that, to realize results, progress in procurement and equipment would require a longer-term advocacy strategy more dependent on the system as a whole and outside of the control of the facility itself. In the future, working with facilities to better identify actionable areas of the local system in these domains, and supporting sites with advocacy approaches for local financing and community participation with in-kind donations, may be an area for improvement in the COPE for CS exercise and action plan implementation.

**The COPE for CS intervention is associated with improvements in facility CS performance**

Facility performance improved as a whole across CS dimensions between baseline and endline. Given the nature of the intervention, these results suggest that facilities successfully advocated for resources and other inputs needed to improve performance, perhaps including trainings or additional staff assigned to the facility. The data show large improvements in the technical capacity of facility staff – especially in health centers. Additionally, health centers developed staff capacity in IUDs,
which goes above and beyond the minimum service requirements for that type of facility. Of note, the projects did not provide trainings or other support to facilities aside from the initial COPE for CS exercise and a brief follow up visit a few months later to discuss action plan progress. Therefore, the changes observed were driven from within each facility.

Additionally, the findings suggest that after participation in the COPE for CS intervention, facilities were better able to accommodate a wide variety of client needs. By endline a large percentage of facilities met requirements to serve clients selecting a new method, or switching from or discontinuing a method requiring removal, as discussed in Table 6. Also, the decreases observed in the number of emergency orders placed may indicate that the CS system within facilities improved as a whole.

**Higher levels of staff commitment to the intervention appears to be associated with greater CS improvements**

More frequent staff discussion of action plans—a proxy for staff commitment—is associated with improved outputs and outcomes, in particular with action item completion. As staff commitment and engagement to facility quality improvement is a key underpinning of the COPE for CS methodology, this finding provides a proof of concept that staff action is a key mechanism of action. Interestingly, we found no evidence for association between the number of official COPE for CS committee meetings and output/outcomes. However, any association between committee meetings and performance improvement is difficult to assess as there was little variation among facilities in the number of these meetings and more data on meeting frequency and timing is needed.

The improvements observed tend to be sustained two years after the initial COPE for CS exercise in the 18 preliminary facilities

This is an important finding regarding sustainability of the intervention. The sub-analysis of the baseline, midline, and endline using data from the preliminary facilities suggests that improvements tend to be sustained over time (measured up to two years post-initiation), despite there being very limited additional investment by the projects. Of note, the preliminary facilities did hold another project-supported facility exercise at the same time that the scale-up sites held their initial exercises. However, the data do not suggest that holding another formal exercise had a major impact on facility performance. Figure 4 shows that preliminary sites achieved their largest gains between 2014 and 2015, with minimal gains after that point.

An interesting observation in assessing the longer-term data is that there appears to be a slight lag in the improvement of equipment scores. While it is impossible to rule out the potential influence of the second supported exercise, it is possible that equipment performance takes more time to improve. Changes in these dimensions may require larger investment, action at higher-level of the health system that may be difficult for a facility to influence, and/or longer-term advocacy and planning. Future research could examine sustainability of the intervention in a more rigorous way, and explore whether there is a need for additional supported facility exercises.

**Limitations**

While this study offers both important data and programmatic reflection on implementation of the COPE for CS intervention, there are several limitations worth noting. First, this study was designed within a programmatic setting where implementation decisions were made in conjunction with local priorities and realities, and not just from a research perspective. Within this context, study districts and facilities received varying levels of support for health services, including FP, from other multilateral agencies and partners. As a result, we cannot assess attribution or make any causal claims. As there is no control group and assignment to the intervention was non-random, a variety of influences not related to the intervention itself could have led to secular changes in performance. Additionally, given the relatively small sample size of 60 facilities in 10 regions, there is not sufficient power to analyze the data using more robust statistical methods; as a result, we were limited to using descriptive analysis and simple regression. We did not adjust for multiple comparisons given the small sample size, the number of comparisons made, and the exploratory nature of this study; a decision that is supported in the statistical literature.  Finally, the implementation period was rather short, and there may not have been adequate time for the output and outcome measures used in this analysis to register change. However, despite these limitations, there appears to be an association between some of the intervention components (in particular, staff commitment) and key outputs/outcomes along hypothesized mechanisms, which lends some support to the intervention being at least partially responsible for the results. Given this, the results of this study warrant more rigorous, future evaluation of this intervention to assess causality.

While several important variables relating to facility-level characteristics are included in the analysis, there remains the potential for unmeasured variables (such as those identified in Figure 1) to confound the relationship between the action plan and the performance score. While stratification according to baseline performance offers important insight on how the intervention may influence performance in low-performing facilities, we do not adjust for these variables. Additionally, it is possible that the results are influenced by regression to the mean in low-performing facilities, we do not adjust for these variables. Additionally, it is possible that the results are influenced by regression to the mean between the baseline and endline measurements (i.e. the tendency of outliers to revert back to mean levels of a variable over time). However, given that the results are consistent according to all performance measures, it does not appear that regression to the mean is the primary driver of the observed results.

In several cases, we constructed key variables of interest ourselves as there was no gold standard measurement available. However, in these cases, we were as rigorous and transparent as possible. For example, as measuring quality of the action plans is inherently somewhat subjective, we developed our measure a priori, so as not to be influenced by the content of action plans, and also used two coders and assessed inter-rater reliability to ensure that the measure was consistent. Additionally, in order to ensure content validity in our measure of CS, we drew heavily...
on the existing literature, standardized indicators, and expert
option to develop the composite scores that we used. While
we attempted to use standardized indicators as much as possible,
some of the standardized CS indicators were not published until
after baseline study implementation; thus, we adapted these
measures as best as possible to the available data. Despite these
ttempts to develop a robust measure, there remains the possi-
ability that the measures do not completely represent all domains
of facility CS performance. Future work in developing and
validating ways to measure performance across CS dimensions
at the facility-level would be a great contribution to the field.

Finally, an important component of the hypothesized mechanism
of action underlying the COPE for CS intervention is engag-
ing the broader community and local leaders in the implementa-
tion of the action plans. The facility assessments did not include
questions on community engagement; however, a separate
analysis of key informant interviews collected from stakehold-
ers participating in the COPE for CS intervention indicate that
strong community engagement through advisory committees was
essential to the success of the intervention. As this is consid-
ered to be an important aspect of the success of the intervention,
community engagement should be measured during future
implementation of COPE for CS.

Conclusions
The results of this study suggest that COPE for CS may be
an effective intervention to improve contraceptive security
at last mile health facilities. Given the dearth of research and
programmatic experience in this area, this study provides
important preliminary programmatic experience, research
insights, and lessons learned upon which future research and
programs attempting to address this important, but often under-
prioritized area of contraceptive security, can build.

Data availability
Underlying data
Harvard Dataverse: Replication Data for: COPE for Contracep-
tive Security Facility Dataset v 2.0, https://doi.org/10.7910/DVN/
7XQL6X

This project includes the following underlying data files:
• Malawi COPE for CS_baseline database_April 2015_
  FINAL stata.xls
• Malawi COPE for CS_ENDLINE database_Feb 2016_
  stata.xls

Extended data
Harvard Dataverse: Replication Data for: COPE for Contracep-
tive Security Facility Dataset v 2.0, https://doi.org/10.7910/DVN/
7XQL6X

This project includes the following extended data files:
• COPE for Contraceptive Security in Malawi Facility
  Assessment Questionnaire
• COPE for Contraceptive Security in Malawi Facility
  Assessment Questionnaire ENDLINE
• Supplemental File 3: COPE for CS action plan quality
  rubric
• Supplemental File 4: Malawi COPE for Contraceptive
  Security: Calculation of Facility CS Performance Scores
• Supplementary Figure 1.

All data are available under the terms of the Creative Commons
Zero “No rights reserved” data waiver (CC0 1.0 Public domain
dedication).

Grant information
The RESPOND project was made possible by the generous
support of the American People through the U.S. Agency for
International Development (USAID), under the terms of the
cooperative agreement GPO-A-000-08-00007-00. Support for
the Malawi Contraceptive Security Project was provided by
PATH, on behalf of the Reproductive Health Supplies Coalition
(NOD.1841-705853-GRT). Additional support for data analysis
was provided by the Bill & Melinda Gates Foundation
(OPP1084384). The views expressed here do not necessarily
reflect those of USAID or the United States Government, the
RHSC or PATH, nor the Bill & Melinda Gates Foundation.

The funders had no role in study design, data collection and
analysis, decision to publish, or preparation of the manuscript.

Acknowledgements
The authors thank Kate O’Connell for reviewing an earlier draft
of this manuscript, Caitlin Shannon for her guidance and review of
preliminary analyses, and Madison Albright for analysis
support. We thank Siti Kachingwe for managing study imple-
mentation and Barwani Msiska and Tiyamike Kamwana for their
assistance. We also thank the Malawi Ministry of Health,
Directorate for Reproductive Health and study facilities for
our partnership, and the Family Planning Association of Malawi
for administrative support.

Author information
Megan E. Schmitt was an employee of EngenderHealth, Inc.
at the time of design, implementation, data collection, analysis
and drafting of this manuscript. She is currently at Public
Health Solutions, New York, USA.
References

33. Staia Statistical Software: Release 14, [computer program], College Station, TX: StataCorp LP, 2015. Reference Source