RESEARCH ARTICLE

Application of behavioral economics principles to reduce injectable contraceptive discontinuation in rural Ethiopia: A stratified-pair, cluster-randomized field trial [version 1; peer review: 2 approved with reservations]

Ali Mehryar Karim 1, Dana Guichon 2, Bantalem Yeshanew Yihun 3, Nebreed Fesseha Zemichael 3, Karina Lorenzana 4, Jeremy Barofsky 2, Wuleta Betemariam 3

1 Global Development, Bill & Melinda Gates Foundation, Addis Ababa, Ethiopia
2 ideas42, New York, New York, 10004, USA
3 The Last Ten Kilometers Project 2020, JSI Research & Training Institute, Inc., Addis Ababa, Ethiopia
4 ideas42, Washington, DC, 20005, USA

Abstract

Background: Contraceptive prevalence in Ethiopia jumped from 6% in 2000 to 36% in 2016, mainly due to increased injectable method use. However, discontinuation rates among injectable users were high (38%). Given that the public sector is the major source for injectable contraceptives, John Snow Inc. (JSI) in collaboration with ideas42 worked with Ethiopia’s flagship Health Extension Program to apply behavioral economics informed approaches to mitigate discontinuation of injectable contraceptives.

Methods: Behavioral diagnosis was conducted to identify the primary drivers of discontinuation of injectable contraceptives. Using these insights, a user-centered behavioral design approach was implemented to create an intervention package, consisting of a health worker planning calendar, a client counseling job aid, and client appointment cards. The study area included two districts from the four regions where JSI was implementing a family planning program. One district from each region was randomly allocated to the intervention arm. Women visiting health posts to use injectable contraceptives were enrolled in the study. Regression methods adjusted for study design, participants’ backgrounds, and contextual factors, were used to estimate the intervention’s effect on discontinuation rates.

Results: Intervention adherence was high for the appointment cards and counseling job aid, but not for the planning calendar. This was not surprising as using appointment cards and the job aid was within the routine workflow of health extension workers, but using the planner was not. The injectable discontinuation rate was 10.8% points lower in the intervention area compared to the control area during the post intervention follow-up.
Conclusion: The use of two behavioral economics informed tools—the appointment card and counseling job aid—effectively decreased injectable discontinuation even with the presence of other health system bottlenecks. Behavioral economics approaches have the potential to enhance family planning programs in Ethiopia and elsewhere.

Trial registration: ISRCTN ISRCTN17390653 (10/04/2019)

Keywords
Behavioral economics intervention, injectable contraceptive discontinuation, family planning

This article is included in the International Conference on Family Planning gateway.

Corresponding author: Ali Mehryar Karim (Ali.Karim@gatesfoundation.org)

Author roles: Karim AM: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Software, Validation, Writing – Original Draft Preparation; Guichon D: Conceptualization, Funding Acquisition, Investigation, Resources, Supervision, Validation, Writing – Review & Editing; Yihun BY: Data Curation, Investigation, Project Administration, Resources, Writing – Review & Editing; Zemichael NF: Investigation, Project Administration, Resources, Supervision, Writing – Review & Editing; Lorenzana K: Conceptualization, Funding Acquisition, Investigation, Resources, Writing – Review & Editing; Barofsky J: Validation, Writing – Review & Editing; Betemariam W: Funding Acquisition, Project Administration, Resources, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by the Bill and Melinda Gates Foundation [OPP1131042]. This work was also supported by the William and Flora Hewlett Foundation [2015-2551].

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

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

How to cite this article: Karim AM, Guichon D, Yihun BY et al. Application of behavioral economics principles to reduce injectable contraceptive discontinuation in rural Ethiopia: A stratified-pair, cluster-randomized field trial [version 1; peer review: 2 approved with reservations] Gates Open Research 2019, 3:1494 (https://doi.org/10.12688/gatesopenres.12987.1)

First published: 13 Jun 2019, 3:1494 (https://doi.org/10.12688/gatesopenres.12987.1)
**Abbreviations**

BE  Behavioral Economics  
CPR  Contraceptive Prevalence Rate  
CBDDM  Community-Based Data for Decision Making  
DHS  Demographic and Health Survey  
FMOH  Federal Ministry of Health of the Government of Ethiopia  
FP  Family Planning  
HEW  Health Extension Worker  
HEP  Health Extension Program  
IUCD  Intra Uterine Cervical Device  
JSI  John Snow, Incorporated  
L10K  Last Ten Kilometers  
LAFP  Long Acting Family Planning  
MNCCH  Maternal, Newborn and Child Health  
PHCU  Primary Health Care Unit  
RMNCH  Reproductive, Maternal, Newborn and Child Health  
SNNP  Southern Nations, Nationalities and Peoples’  
WDA  Women Development Army

**Introduction**

The national family planning program in Ethiopia has made considerable strides in providing access to services across the country. The contraceptive prevalence rate (CPR) in Ethiopia jumped from 6% in 2000 to 36% in 2016. This increase was mainly the result of a sharp increase in the use of injectable contraceptive methods, from 3% in 2000 to 23% in 2016. Nonetheless, 35% of contraceptive users discontinued use within 12 months of uptake, while the discontinuation rate among injectable contraceptive users was 38% (Central Statistical Agency and ICF, 2016).

The Ethiopian Federal Ministry of Health (FMOH) aims to increase the CPR to 55% by 2020 (Federal Ministry of Health, 2015). High rates of discontinuation pose a threat to the achievement of this target, as well as to the intentions of many Ethiopian women to space births and to limit the number of children they have. High rates of contraceptive discontinuation impose a cost on the woman, her family, and her country’s health system. Discontinuation can lead to unintended pregnancies, causing social, economic, and emotional distress, and also contributing to maternal morbidity and mortality (Ahmed et al., 2012; Ali et al., 2012; Cleland et al., 2012).

Women may choose to discontinue a contraceptive method because their reproductive intentions have changed, and they no longer intend to use family planning. However, if women are discontinuing, but still have intentions to use contraception, the system has not met their needs. Discontinuation rates may be high for a variety of reasons. Some of these are problems at the system level—for example, high rates of stockouts or distance from the source of the family planning method could be structural barriers impeding continued use. However, various nonstructural factors may also contribute. Often, the dominant reason for stopping a modern, reversible method is dissatisfaction (World Health Organization, 2012). Dissatisfaction with the method could stem from gaps in counseling that lead to suboptimal initial method choices, health concerns or side effects, or choice of an inconvenient method, among other reasons (Bradley et al., 2009). In contexts where structural barriers have for the most part been eliminated, nonstructural, or behavioral, factors may be driving discontinuation. Ethiopia is likely to be such a context. The introduction of the country’s flagship Health Extension Program (HEP) has brought family planning services to the doorsteps of the rural population (Admasu et al., 2016).

Behavioral economics (BE) offers a promising toolkit for designing interventions to change health-related behaviors and decision-making, from both the provider and the client perspective, by improving understanding of why people behave as they do and what motivates their decision-making and action (Buttenheim & Asch, 2013; Datta & Mullainathan, 2014). Insights from BE and a structured behavioral design methodology have recently been used to improve reproductive health programming (Spring et al., 2016), suggesting that the approach could be applied to mitigate the contraceptive discontinuation problem in Ethiopia.

Abandonment of injectable methods is an urgent matter, given that they are the primary method of choice among married women of reproductive age in Ethiopia, and that the government sector provides 82% of contraceptive methods (Central Statistical Agency and ICF, 2016). The Last Ten Kilometers (L10K) Project of John Snow Research & Training Institute, Inc. (JSI) collaborated with ideas42 to work with the HEP to design and test behavioral approaches to mitigate the problem of discontinuation of injectable contraceptives.

The overall goal of this study is to increase the rate of contraceptive continuation among women of reproductive age who are new users of injectable contraceptives in rural Ethiopia. The primary objectives in support of this goal are 1) to use behavioral insights to develop a package of interventions that will improve the rate of contraceptive continuation of women using injectable contraceptives in Ethiopia and 2) to trial this package of interventions and assess the preliminary impact of this package and its limitations.

**Methods**

**Study context**

The health care delivery system of rural districts in Ethiopia is comprised of three to four primary health care units (PHCUs) supported by one primary hospital. A PHCU is formed by one health center for every 25,000 people in the district, with five satellite health posts (Federal Ministry of Health, 2015). Health centers are staffed by health officers, nurses, midwives, and laboratory technicians, each with an undergraduate diploma or health degree. They provide preventive and curative services, including basic emergency obstetric and newborn care, as well as supervision of satellite health posts.
The health posts employ two female Health Extension Workers (HEWs) serving a community (kebele) of about 5,000 people with basic community-based promotive, preventive, and curative health services. This system forms the nucleus of the flagship HEP in Ethiopia (Wakabi, 2008). A network of Women’s Development Army (WDA) volunteers supports the HEWs to ensure community engagement in the delivery of HEP services.

The WDA network is organized into groups of 30 households led by one WDA team leader, with subgroups of six households led by one WDA member (Admasu et al., 2016).

The health posts provide condoms, oral contraceptive pills, emergency contraception, injectable contraceptives, and Implanon insertion services; health centers offer intrauterine contraceptive devices (IUCDs) and implant insertion and removal services in addition to the methods provided at health posts. Permanent methods are offered at primary hospitals or higher-level facilities.

The L10K program sought innovative, community-based strategies for the HEP to provide high-impact reproductive, maternal, newborn, and child health (RMNCH) services. To this end, L10K partnered with 10 local civil society organizations to enhance the interactions between HEP frontline health workers (HEWs and WDAs) and communities to achieve more accessible, efficient, and equitable maternal, newborn, and child health (MNCH) services (Darmstad et al., 2013). The L10K intervention area covered 115 districts in the Amhara, Oromia, Southern Nations, Nationalities and Peoples’ (SNNP), and Tigray regions, reaching a population of 17 million people, approximately 19% of Ethiopia’s population.

In eight of its 115 intervention districts (two districts per region), L10K collaborated with the respective Regional Health Bureaus to implement interventions to improve the demand for and quality of community-based long-acting family planning (LAFP) services. The eight districts were purposely selected by the respective Regional Health Bureaus to avoid duplication of other partners’ work. The family planning interventions began in 2014, and were implemented in one of the four PHCUs within each of the eight selected districts. Intervention activities included family planning counseling training for HEWs, comprehensive family planning for the health center service providers, supplies of national tools and job aids for HEWs, institution of quarterly review meetings, and enhancement of referral and linkages with the health center and health posts for LAFP and postpartum family planning services. The eight PHCU where L10K was implementing the FP interventions were selected for participation in the BE study.

The intervention
To identify behavioral barriers impeding continued use of injectable contraceptives, design a targeted solution, and test the impact of this approach, the study team applied a four-stage behavioral diagnosis and design methodology. This process involved four distinct phases depicted in Figure 1: problem definition, behavioral diagnosis, intervention design, and testing (Datta & Mullainathan, 2014).

**Problem definition (Step 1).** The general problem of contraceptive discontinuation is broad, and it encapsulates diverse methods, subpopulations, and manifestations of discontinuation. Thus, the goal of the problem definition process was to (a) identify the different problem definition candidates that represent these multiple dimensions and components of discontinuation in rural Ethiopia, and (b) evaluate these candidate definitions systematically along a set of criteria so that the strongest and most actionable problem or problems are identified. The criteria were:

1. The key drivers for discontinuation of contraceptives included behavioral issues among others (e.g., cultural issues, problems with the supply chain).
2. The existence of an intention-action gap, i.e., women have the desire to continue using contraceptives, but fail to follow through on this intention and discontinue.
3. The problem could be moved by a behavioral intervention; and,
4. Structures were in place for implementing clustered-randomized clinical trials to test the effectiveness of the behavioral intervention.

Problem definition included the analysis of the 2011 Ethiopian Demographic and Health Survey, informal conversations with L10K senior and mid-level managers, process mapping, preliminary behavioral mapping, and a field visit in October 2014. The field visit included conversations with frontline HEP workers (i.e., HEWs and the WDA) and contraceptive users, along with observation of family planning provision.

The process gathered anecdotal evidence that both passive (unconscious decisions) and active (conscious decisions) contraceptive discontinuation could be occurring. All long-acting (i.e., IUCD and implants) and short-acting (i.e., injectables, pills, and condoms) reversible methods may be subject to active discontinuation, while only short-acting reversible contraceptive methods may be subject to passive discontinuation.

**Figure 1. Stages of Behavioral Design methodology (Datta & Mullainathan, 2014).**
Several candidates for defining the problem were identified and evaluated based on the four criteria described above. The final problem definition from the process was: “Women of reproductive age who are using injectable contraceptives discontinue use within 12 months of uptake.”

**Behavioral diagnosis (Step 2).** The objective of the behavioral diagnosis was to identify the primary behavioral drivers or the “behavioral bottlenecks” leading to discontinuation of injectable contraceptives within the study population. The process included detailed behavioral mapping to generate hypothesized bottlenecks; contextual reconnaissance to gather evidence; and synthesis of evidence and evaluation of hypothesized bottlenecks. Evidence gathered during the contextual reconnaissance stage included data from client and provider interviews, observations of family planning provision, analysis of data from Demographic and Health Surveys, conversations with project staff, and analysis of secondary literature. This evidence was used to confirm, reject, or refine the hypothesized bottlenecks.

The set of hypotheses that seemed to describe the primary bottlenecks driving observed behavior (i.e., discontinuation of injectable contraceptives) and its intervention design implications are presented in three categories: bottlenecks contributing to passive discontinuation, active discontinuation, and provider behavior (Table 1).

**Intervention design (Step 3).** Building directly on the behavioral diagnosis and related design implications, a package of scalable interventions was selected and refined to ensure that the interventions were feasible, culturally acceptable, aligned with HEP policies, contextually appropriate, and human-centered. To ensure human centered design, the tools were co-created with the HEWs and clients. The intervention design package included three components—a HEW planning tool located at the health post; a client counseling job aid used by HEWs during service provision; and appointment cards distributed by HEWs during services provision and used by clients (Figure 2).

**HEW planning calendar:** The planning calendar is a desktop-based calendar, located at the health post and utilized by the HEW. The planning calendar serves as an aid for the HEW when scheduling future appointments and outreach, and it guides her toward better time management and planning.

**Client counseling job aid:** The counseling job aid is a simple guide to frame the HEW service provision interaction with clients to (1) prompt HEWs to provide high-quality services at every interaction and to create an atmosphere in which the client feels comfortable and supported, (2) prompt the key components of family planning counseling, and (3) reduce the extent to which a provider needs to remember exact protocols for setting return appointments and managing records. Three job aids—one for counseling on available family planning methods for informed choice; one for counseling new family planning clients; and one for counseling returning clients—are displayed on three sides of a triangular tower and kept on the HEW’s counseling table.

**Appointment cards:** HEWs distribute appointment cards to all clients requiring return appointments. They remind clients about future appointments and provide critical information on what to do if they cannot make their appointments or experience health concerns in the interim.

**Test (Step 4).** The final step of the behavioral design methodology was to test the package of interventions to assess implementation feasibility and impact on the outcomes of interest.

**Trial design**
A stratified-pair cluster-randomized field trial was implemented to test the package of BE interventions. The stratification variable was the administrative region. One of the two FP intervention PHCUs (i.e., districts) from each region was randomly selected for the intervention arm, while the other served as the control arm. All 19 satellite health posts in the four intervention PHCUs received the BE intervention, while the 21 health posts in the four PHCUs of the control arm did not receive it (Table 2). In Adinebried, Alefa, and Ebot Tiora PHCUs, the BE intervention package was initiated in February 2016; and in Sentema began in March 2016. Masking the intervention package was not possible.

**Trial implementation**
Two-day training was developed, covering the purpose of the study and the implementation of the intervention package. All HEWs including their supervisors at the District Health Office and health centers were trained. A family planning client register was introduced at the health posts so that study participants could be traced and interviewed to estimate injectable contraceptive discontinuation rate. The HEWs including their supervisors in the control districts were trained on to use the family planning client register.

L10K staff conducted routine supportive supervision of the study area mainly to support its broader family planning intervention.

**Study participants**
The study participants were women who visited the study area health posts during the intervention period and fulfilled any of the following three criteria: (1) initiating use of injectable contraceptives; (2) returning to the use of injectable contraceptives after at least a six-month lapse; or (3) switching to injectable contraceptives from another family planning method. Those who did not meet any of these three criteria was excluded from the study.

**Outcome measurements**
The primary outcome of interest was injectable contraceptive discontinuation rate within 12 months of uptake. To measure it, first the duration of injectable contraceptive use was measured from the three survey items: (1) the month and year the eligible participant visited the health post between intervention
Table 1. Overview of behavioral bottlenecks and their design implications.

<table>
<thead>
<tr>
<th>Behavioral Bottlenecks</th>
<th>Design Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bottlenecks Contributing to Passive Discontinuation</strong></td>
<td></td>
</tr>
<tr>
<td>1. Prospective Memory Failure to Follow Through on Next Injection: Due to prospective memory failure, women do not come to the health post at the appropriate time.</td>
<td>Salient, timely reminders for women.</td>
</tr>
<tr>
<td>2. Tunneling Leads to Myopic Planning or Failures to Plan for Next Injection: Women tunnel on the specific date of a follow-up appointment, yet they do not form any contingency plans for receiving their injection if they cannot make that exact date.</td>
<td>Prompt plan-making and contingency planning for return appointments.</td>
</tr>
<tr>
<td>3. Hassle Factors: Women may be deterred by hassles associated with successful continuation of injectable contraceptives.</td>
<td>Timely reminders of fertility intentions.</td>
</tr>
<tr>
<td><strong>Bottlenecks Contributing to Active Discontinuation</strong></td>
<td></td>
</tr>
<tr>
<td>4. Perception of Limited Choice Set for Continued Use: Women experience and wish to eliminate side effects from injectable contraceptives, yet they do not feel that they have viable alternatives, nor do they consider switching methods.</td>
<td>Expand the choice set of methods women consider to be viable options; increase and maximize HEW and WDA touch-points with women.</td>
</tr>
<tr>
<td><strong>Bottlenecks Contributing to Provider Behavior</strong></td>
<td></td>
</tr>
<tr>
<td>5. Bandwidth Tax: HEWs have limited time and attention to devote to client tracking and case management. Thus, the date of clients’ follow-up appointment is not salient to HEWs at the appropriate time to provide prospective appointment reminders and retrospective follow-up for missed appointments.</td>
<td>Simplify case management systems to reduce cognitive load of HEWs.</td>
</tr>
</tbody>
</table>

![Appointment Card](image1.png) ![Client Counseling Job Aid](image2.png) ![HEW Planning Tool](image3.png)

Figure 2. Clockwise from left: appointment card, client counseling job aid and Health Extension Worker (HEW) planning calendar.

onset and November 2016 (uptake month); (2) whether the participant was still using the injectable contraceptive without interruption at the time of the survey, or with a maximum of a month (30 days) of interruption (survey month); and (3) if the participant was not currently using the injectable contraception, then the month and year the participant had last obtained injectable contraceptive without interruption or with a maximum of one month of interruption (dropout month). The duration of injectable contraceptive use from uptake is the difference between uptake month and survey month, or
uptake month and dropout month. The duration was measured in months. In a few of the dropout cases, where the participant did not accurately describe the last time she obtained an injectable contraceptive (without interruption), then the duration of use was obtained by asking how long she had used injectable contraception without interruption.

The intermediate outcomes considered were mean number of correct side effects of injectable contraceptive spontaneously recalled by the participant; mean number of different types of contraceptive methods spontaneously recalled by the provider; mean number of correct actions to be taken spontaneously recalled by the participant if she experienced side effects from injectable contraceptives; whether the participant was told about the side effects of injectable contraceptives by the provider; whether the participant was told what to do if she experienced side effects; whether the participant was informed about other methods of contraception; whether the participant was told, during her last visit to the health post for injectables, about a local holiday or event as a reminder of the next appointment date; whether the participant was given an appointment card during the last visit to obtain an injectable contraceptive; and whether the participant missed an appointment for injectable contraceptives.

Sample size
The sample size for the study depended on the caseload of the health posts and the duration of the recruitment period. Initially, it was decided that the recruitment period for the study would be six months. Based on the caseload records of the health post, it was estimated that in six months about 660 eligible participants (330 from each study arm) could be recruited. If the expected discontinuation rate of injectable contraceptives within 12 months of uptake without the BE intervention was at 34% (the level reported in the Ethiopian Demographic and Health Survey conducted in 2011) (Central Statistical Agency and ICF, 2012), the sample size that would be obtained in six months would have 80% power to detect at least a 14 percent-age-point reduction in injectable discontinuation rate with a 95% confidence level, if the cluster-randomization study design effect was 2.0 and the nonresponse rate was 5%.

Data collection
Data collection was planned to take place 12 months after September 2016, i.e., in October 2017. However, due to competing priorities, the data collection was conducted in December 2017. Eligible participants were drawn from women of reproductive age who visited the study area health posts for injectable contraceptive services between the onset of the intervention and November 2016.

Ethical clearance for the study was obtained from each of the four Regional Health Bureaus and from the JSI Ethical Clearance Committee. Two types of questionnaires were pretested and used—one for the health post in the study area and the other for the study participants (see extended data (Karim, 2019)). Data collection from the health posts was done in October 2017 to assess status of the BE interventions and to assess the status of the overall health system functions using a structured questionnaire. One day of training for data collection from the health posts was provided to four public health professionals and who had previously conducted health facility surveys for L10K. The supervisors of the health post data collection were L10K regional office staff. Survey responses were electronically captured using SurveyCTO, a web-based survey data-collection platform using mobile technology (Dobility Inc., 2018).

The study participant questionnaire was pre-coded and translated into three major local languages: Amharic, Oromia, and Tigrigna. In SNNP, where there are several ethnic languages, the interviewers translated questions from Amharic informally. In total, 28 public health professionals were given a two-day training on collecting data from the study participants, which included practical exercises. The data collectors had previously conducted household surveys for L10K. Eight of the data collectors who had good survey experience from conducting previous L10K surveys were oriented to supervise the data collection. The data collectors did not know which was the intervention arm and which was the control arm. The list of women of reproductive age who visited the study area health posts for injectable contraceptive services between the onset of the BE intervention and November 2016 was obtained from the

Table 2. List of primary health care units (PHCUs) by study arm. SNNP - Southern Nations, Nationalities and Peoples’.

<table>
<thead>
<tr>
<th>Study Arm</th>
<th>Region</th>
<th>District</th>
<th>PHCU</th>
<th># of Health Posts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>Amhara</td>
<td>Ensaro</td>
<td>Lemi</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Oromia</td>
<td>Chora</td>
<td>Kumbabe</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>SNNP</td>
<td>Yem</td>
<td>Fofa</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Tigray</td>
<td>Samre Sehart</td>
<td>Finarwa</td>
<td>3</td>
</tr>
<tr>
<td>Intervention</td>
<td>Amhara</td>
<td>Burie Zuria</td>
<td>Alefa</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Oromia</td>
<td>Seka-Chekorssa</td>
<td>Sentema</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>SNNP</td>
<td>Dalocha</td>
<td>Ebot Tirora</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Tigray</td>
<td>Laelay Adiabo</td>
<td>Adinebried</td>
<td>5</td>
</tr>
</tbody>
</table>
family planning client registers. All the women in the register were visited at their households, and those who met the eligibility criteria and consented to participate were interviewed. The data collection from the study participants was done in December 2017. Survey responses were electronically captured using SurveyCTO.

**Independent variables.** The major independent variable of interest is an indicator variable for the intervention area (study arm). The effect of the intervention was controlled for background characteristics of the participants and the health system functioning status indicators of the health posts. The background characteristics considered were age, education, number of children, age of last child, and wealth quintile. A wealth index score was constructed for each participant’s household using principal component analysis of household possessions (electricity, watch, radio, television, mobile phone, telephone, refrigerator, table, chair, bed, electric stove, and kerosene lamp), and household characteristics (roof material, wall material, type of latrine, and source of drinking water). Households were ranked according to the wealth score and then divided into five quintiles (Filmer & Pritchett, 2001).

The health system functioning status indicators measured for the study area health posts were population to HEW ratio of the kebele; household to active WDA team leader ratio of the kebele; number of Level 3 and Level 4 HEWs at the health post; whether the health post was closed for a month or more during the observation period; availability of contraceptives at the health post during the day of the visit; frequency of supportive supervision visits received by the health post; whether the health post maintained and updated the health management information system records; whether the health post maintained a tickler file system for family planning clients; and whether the health post had updated records for activities related to the Community-Based Data for Decision Making (CBDDM) strategy of the L10K project. Adherence to the BE intervention was measured by observing the health post for the presence of appointment cards, placement of counseling job aid on the counseling table, and use of HEWs’ planning calendar.

**Statistical analysis**
All statistical analysis was done using Stata version 15.1 (StataCorp, 2017). The availability and use of BE tools in the intervention area health posts was assessed using descriptive statistics. Then, the distributions of study participants’ background characteristics were analyzed and statistically significant differences in the background characteristics between the study arms were assessed using Pearson’s chi-squared statistics, adjusted for clustering of the respondents within the health posts.

Distributions of the study participants according to health post characteristics were compared between study arms, and the statistically significant differences were assessed, also using adjusted Pearson’s chi-squared statistics. The characteristics of the health posts were the health system functioning status indicators described in the measurement section above.

The intervention effects on the intermediate outcomes were based on intention to treat analysis. Logit regression was used when the outcome variable was binary, and ordinary least square regression when the outcome variable was continuous. The analysis accounted for the stratified PHCUs, and all the independent variables described in the measurement section above. Stepwise backward selection was implemented to identify regressors for each of the outcome models with p-values set at 0.2 for variables to enter or exit the model. The study arm indicator and the indicators for the matched PHCUs were forced to remain in the model. For each outcome, a health post-level random effects regression model was estimated using the independent variables from the backward selection process. If the health post level random effect was statistically significant (p < 0.05), then the health post-level random effects model was the final model; and if not, then the model identified by the backward selection was the final one. The goodness-of-fit of the final model was assessed using a global F-test, global Wald’s statistic, or global likelihood ratio test, depending on the estimation method for the final model. The final model was used to obtain adjusted point estimates of the outcome and its 95% confidence intervals according to the study arm. The adjusted differences in the outcomes between the two study arms and its 95% confidence intervals were also estimated.

The intervention effect on the primary outcome was estimated with a discrete hazard model using logistic regression (Allison, 1982; Canette, 2016). Data were organized and analyzed using the methodology described by Canette (2016). The discrete units were months. The identification of the final model was done using the same procedure that was used to identify the models for the intermediate outcomes described above. The final model was used to estimate the injectable discontinuation rate during the 12 months following uptake.

**Results**
In general, the catchment populations of the health posts in the intervention areas were nearly twice those of the control areas. The average catchment population of the health posts in the intervention area was 7,868; for the control area it was 4,064.

Assessment of the implementation status of BE tools at the health posts indicates that 12 of the 19 health posts in the intervention area had appointment cards available on the day of the visit, while only two health posts were stocked out of the appointment cards for more than a year (Table 3). Client counseling towers were available in all the intervention area health posts, and they were on top of each HEW’s counseling table. However, the HEWs from one of the intervention area health posts indicated that they did not use the counseling job aid. Use of HEW the planning calendar was poor. Only four of the intervention area health posts were using the planning calendar appropriately, while the rest were not using it during the four weeks preceding the survey, or they did not have the calendar for the most current year.

Health post records indicated that 2,490 women in the intervention area and 1,684 women in the control area visited the
study area health posts between February 2016 and November 2016 to obtain injectable contraceptives (Table 4). Among them, 335 women (13%) from the intervention area and 408 women (24%) from the control area were eligible for the study and were interviewed after obtaining consent; 688 women (28%) in the intervention area and 559 women (33%) in the control area could not be located due to erroneous records; while 420 women (17%) in the intervention area and 167 women (10%) in the control area could not be reached for interview. If the eligibility rate for participating in the study among those who could not be located for an interview was similar to those who were located, then the loss to follow-up of eligible study participants was 22% in the intervention area and 18% in the control area. There were no non-responses among those who were eligible.

Table 5 compares the background characteristics of the study participants between the intervention and control areas. Compared to those of the control area, the intervention area participants were less likely to have higher education ($p < 0.05$); and the age of the last child was less likely to be more than three years ($p < 0.001$).

Table 6 shows the distribution of the study participants according to the health system functional status of the health posts to which they belonged. Although it appeared that compared to participants from the control areas, participants of the intervention areas were more likely to be in the catchment area of health posts with better-functioning health systems (i.e., more Level 4 HEWs, higher density of active WDA team leaders, fewer contraceptive stockouts, more frequent supportive supervisory visits, more frequently updated family folders, more updated CBDDM registers, and less likelihood of closure for more than a month during the past two years), the differences were not statistically significant. However, the HEWs in the intervention area were serving a larger population. Of the participants in the intervention area, 81% were in the catchment area of health posts that had more than 2,500 people for every HEW deployed; compared to 37% in the control area ($p < 0.01$).

The frequency distribution of person-months of follow-up according to the independent variables can be found in Table A1 of the data repository (see extended data (Karim, 2019)).

Table 7 gives the adjusted point estimates of the outcomes of interest by study arm and the intervention effects with 95% confidence intervals. The adjusted estimates considered the differences between the study arms observed in Table 5 and Table 6. Figure 3, Figure 4 and Figure 5 display the analysis in Table 7 using figures. The mean number of side effects from injectable contraceptive methods recalled was 0.24 higher ($p < 0.01$) among the intervention area participants than among those in the control areas. The proportion who recalled being told of a local holiday or event as a reminder for the next appointment for injectable contraceptives was 16.7% points higher (95% Confidence Interval [CI]: 7.4%, 26.0%) among the intervention area participants than among those from the control areas ($p < 0.001$). Although the intervention’s effects on the proportion told about side effects, proportion told what to do if they had side effects, the proportion told about other methods, the proportion given appointment card, and the proportion who missed any appointment were in the expected directions, they were not statistically significant ($p > 0.1$).

The primary outcome of interest, i.e., injectable contraceptive discontinuation rate within 12 months of uptake, was 10.8% points lower (95% CI: 2.4%, 19.1%) among intervention area…….
Table 4. Frequency Distribution of family planning clients who visited the health posts in the study area between February 2016 and November 2016.

<table>
<thead>
<tr>
<th>Category</th>
<th>Control # (%)</th>
<th>Intervention # (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number eligible and interviewed</td>
<td>408 (24)</td>
<td>335 (13)</td>
</tr>
<tr>
<td>Number not eligible</td>
<td>550 (33)</td>
<td>688 (28)</td>
</tr>
<tr>
<td>Number could not be located (erroneous records)</td>
<td>559 (33)</td>
<td>1,047 (42)</td>
</tr>
<tr>
<td>Number could not be reached</td>
<td>167 (10)</td>
<td>420 (17)</td>
</tr>
<tr>
<td>Total</td>
<td>1,684 (100)</td>
<td>2,490 (100)</td>
</tr>
</tbody>
</table>

Table 5. Percentage distribution of study participants by background characteristic. SNNP - Southern Nations, Nationalities and Peoples.

<table>
<thead>
<tr>
<th>Sample characteristics</th>
<th>Control (N = 408)</th>
<th>Intervention (N = 335)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16–22</td>
<td>26</td>
<td>26</td>
<td>0.441</td>
</tr>
<tr>
<td>22–29</td>
<td>41</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>30–39</td>
<td>26</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>40–49</td>
<td>7</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td>0.043</td>
</tr>
<tr>
<td>None</td>
<td>56</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>19</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Higher</td>
<td>25</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Age of last child</td>
<td></td>
<td></td>
<td>&lt; .001</td>
</tr>
<tr>
<td>&gt; 3 years</td>
<td>92</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>≤ 3 years</td>
<td>7</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>No children</td>
<td>2</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Religion*</td>
<td></td>
<td></td>
<td>0.480</td>
</tr>
<tr>
<td>Orthodox</td>
<td>60</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>Protestant</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Muslim</td>
<td>38</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>Wealth quintile</td>
<td></td>
<td></td>
<td>0.273</td>
</tr>
<tr>
<td>Lowest</td>
<td>17</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Second</td>
<td>13</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>Middle</td>
<td>23</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Fourth</td>
<td>25</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Highest</td>
<td>22</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Region</td>
<td></td>
<td></td>
<td>0.393</td>
</tr>
<tr>
<td>Amhara</td>
<td>16</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Oromia</td>
<td>41</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>SNNP</td>
<td>23</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Tigray</td>
<td>20</td>
<td>42</td>
<td></td>
</tr>
</tbody>
</table>

*For the statistical test, the Protestants were collapsed with Orthodox Christians.
Table 6. Distribution of study participants according to the health system functional status indicators of the health posts. HEW – Health Extension Workers, WDA – Womens Development Army, CBDDM – Community-Based Data for Decision making.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Control</th>
<th></th>
<th>Intervention</th>
<th></th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% of Participants (N = 408)</td>
<td># of Health Posts (N = 21)</td>
<td>% of Participants (N = 335)</td>
<td># of Health Posts (N = 19)</td>
<td></td>
</tr>
<tr>
<td>Population per HEW in the kebele</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 2,500</td>
<td>63</td>
<td>14</td>
<td>9</td>
<td>2</td>
<td>0.006</td>
</tr>
<tr>
<td>2,501 to 5,000</td>
<td>30</td>
<td>5</td>
<td>84</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>&gt; 5,000</td>
<td>7</td>
<td>1</td>
<td>7</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Number of Level 3 HEWs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.972</td>
</tr>
<tr>
<td>0</td>
<td>22</td>
<td>5</td>
<td>19</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>29</td>
<td>6</td>
<td>29</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>2+</td>
<td>49</td>
<td>10</td>
<td>52</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Number of Level 4 HEW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.564</td>
</tr>
<tr>
<td>0</td>
<td>53</td>
<td>12</td>
<td>34</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>23</td>
<td>4</td>
<td>38</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>5</td>
<td>27</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Households per active WDA team leader</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.204</td>
</tr>
<tr>
<td>&lt; 50</td>
<td>12</td>
<td>3</td>
<td>41</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>50-100</td>
<td>54</td>
<td>10</td>
<td>34</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>&gt; 100</td>
<td>34</td>
<td>8</td>
<td>25</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Number of methods stocked out on day of visit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.647</td>
</tr>
<tr>
<td>0</td>
<td>49</td>
<td>9</td>
<td>65</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>37</td>
<td>9</td>
<td>26</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>2+</td>
<td>14</td>
<td>3</td>
<td>9</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Supportive supervisory visit received</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.164</td>
</tr>
<tr>
<td>Last month</td>
<td>40</td>
<td>7</td>
<td>60</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>In last 3 months</td>
<td>26</td>
<td>6</td>
<td>32</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>&gt; 3 months</td>
<td>35</td>
<td>8</td>
<td>8</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Have family folder</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>100</td>
<td>21</td>
<td>100</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Household demographics of family folders</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.412</td>
</tr>
<tr>
<td>updated in last 12 months</td>
<td>33</td>
<td>7</td>
<td>20</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>67</td>
<td>14</td>
<td>80</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Family planning tickler file</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.761</td>
</tr>
<tr>
<td>Currently used</td>
<td>91</td>
<td>18</td>
<td>87</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Not currently used</td>
<td>3</td>
<td>1</td>
<td>8</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Not available</td>
<td>6</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>CBDDM register updated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.090</td>
</tr>
<tr>
<td>All updated</td>
<td>24</td>
<td>6</td>
<td>62</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Partially updated</td>
<td>17</td>
<td>4</td>
<td>10</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Not available/not updated</td>
<td>59</td>
<td>11</td>
<td>28</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Health post closed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.310</td>
</tr>
<tr>
<td>No</td>
<td>93</td>
<td>19</td>
<td>98</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>For &gt; 1 month</td>
<td>7</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
Table 7. Adjusted effects of the intervention on intermediate and primary outcomes.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Control</th>
<th>Intervention</th>
<th>Difference (95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean # of side effects of injectable contraceptive recalled</td>
<td>1.43</td>
<td>1.67</td>
<td>0.24 (0.07, 0.41)</td>
<td>0.006</td>
</tr>
<tr>
<td>Mean # of contraceptive methods available recalled</td>
<td>1.67</td>
<td>1.67</td>
<td>-0.00 (-0.13, 0.14)</td>
<td>0.962</td>
</tr>
<tr>
<td>Mean # of actions that can be taken for side effects recalled</td>
<td>1.16</td>
<td>1.16</td>
<td>0.00 (-0.09, 0.09)</td>
<td>0.999</td>
</tr>
<tr>
<td>% told about side effects</td>
<td>39.8</td>
<td>51.5</td>
<td>11.7 (3.2, 29.4)</td>
<td>0.021</td>
</tr>
<tr>
<td>% told what to do if they had side effects</td>
<td>48.0</td>
<td>51.5</td>
<td>3.5 (3.1, 8.6)</td>
<td>0.020</td>
</tr>
<tr>
<td>% told about other methods</td>
<td>78.2</td>
<td>80.5</td>
<td>2.3 (1.2, 3.5)</td>
<td>0.001</td>
</tr>
<tr>
<td>% told about a local holiday or event as a reminder for appointment</td>
<td>60.3</td>
<td>77.0</td>
<td>16.7 (11.7, 21.7)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>% given an appointment card</td>
<td>60.6</td>
<td>73.6</td>
<td>12.9 (7.7, 18.1)</td>
<td>0.003</td>
</tr>
<tr>
<td>% missed any appointment for an injectable contraceptive</td>
<td>15.7</td>
<td>11.8</td>
<td>-3.9 (-9.4, 1.7)</td>
<td>0.169</td>
</tr>
<tr>
<td>% discontinued injectable contraceptive within 12 months of uptake</td>
<td>52.6</td>
<td>41.8</td>
<td>-10.8 (-19.1, -2.4)</td>
<td>0.012</td>
</tr>
</tbody>
</table>

Figure 3. Mean number of side effects, family planning (FP) methods and action to take for side-effects recalled by the participants.

Discussion

This is a unique study that applies a behavioral design methodology to reduce injectable contraceptive discontinuation in rural Ethiopia. Following a behavioral diagnosis to understand the behavioral barriers leading injectable contraceptive users to discontinue use, a package of interventions was designed and implemented. Behavioral factors (among others) that contributed to discontinuation of injectable contraceptives were found to be mainly active or passive from the client perspective, or contributed to the providers’ action or lack of action. The hypotheses for passive discontinuation were prospective memory failure to follow through on the next injection, myopic planning or failures to plan for the next injection, and hassle factors. The hypothesis for active discontinuation was the perception of a limited choice set for continued use of contraception. The hypothesis on the provider’s perspective was that the provider had limited time and attention to devote to client tracking and case management. Based on these behavioral insights, a package of interventions was designed and tested. The results of this study indicate that the BE interventions were effective in decreasing injectable contraceptive discontinuation in the study environment, thus supporting the hypotheses generated during the
behavioral diagnosis phase. The analysis of the intervention’s effect controlled for the background characteristics of the participants and for the health system functional status indicators of the health posts.

Erroneous identification records of injectable users in the study area health posts were substantial. It is common knowledge among HEP workers, though not well documented, that some contraceptive users choose to use injectables discreetly and provide false identities to obtain services. The data collectors confirmed this assumption with the HEWs in the study area. Poor record keeping at the health posts may also be partly responsible for the errors.

Loss to follow-up could have biased the study result. However, if the injectable contraceptive discontinuation rate

---

### Figure 4. Percentage of participants who received method information from provider, given appointment card and missed an appointment.

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Told about side-effects</td>
<td>40%</td>
<td>50%</td>
</tr>
<tr>
<td>Told what to do if had side-effects</td>
<td>49%</td>
<td>51%</td>
</tr>
<tr>
<td>Told about other methods</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Told about a local holiday as a</td>
<td>61%</td>
<td>74%</td>
</tr>
<tr>
<td>reminder for next appointment*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Given appointment card</td>
<td>61%</td>
<td>74%</td>
</tr>
<tr>
<td>Missed appointment</td>
<td>16%</td>
<td>12%</td>
</tr>
</tbody>
</table>

*p < .01

---

### Figure 5. Injectable contraceptive discontinuation rate.

*The difference in discontinuation rate was statistically significant (p > .05)*
among women who were lost to follow-up was similar to that of women who were not lost to follow-up within the respective study arms, then the estimate of the intervention effect was unbiased. The possible bias due to loss to follow-up was minimized by the study design and by the application of regression methods.

The primary threat to the validity of the randomization is that there was no baseline on a wide range of characteristics and that some of the treatment effects are highly sensitive to the specific controls used. To validate these treatment effects (especially for the main outcome we care about) we argue that the specific model and controls derived from stepwise backward selection are the correct ones and that unobserved differences do not drive the difference.

The existing family planning client tracking system for HEWs was the tickler filing system to track defaulters (Chewicha & Azim, 2013), which appeared not to be a useful tool for prospective appointment reminders to avoid defaulting appointments for the next dose of injectable contraceptives. The planning calendar was designed to help the HEWs to provide prospective follow-up reminders to the FP clients. However, it is unlikely that the HEW planning calendar had any effect on the outcomes of interest, mainly because nine of 19 intervention area health posts did not have the most recent planning calendar, and among the 10 health posts that had copies of the most recent planning calendar, six were not using them. One of the possible reasons for not using the planning calendar could be that the HEWs were already overwhelmed by the existing recording requirements of the HEP, and the planning calendar was an extra recording burden for them. The planning calendar was a parallel system to the HEWs’ existing record-keeping efforts.

The study participants in both arms reported receiving appointment cards from the HEWs during the last time they visited the health post to obtain injectable contraceptives. This was not surprising, because during the problem definition and behavioral diagnosis exercises it was noticed that some HEWs provided injectable contraceptive clients a paper note as a reminder for the next appointment. Thus, formal appointment cards formalized an already used ad-hoc strategy used by the HEWs as one of the intervention tools. While providing the appointment card and counseling clients, the HEWs in the intervention area were advised to relate the appointment date for the next dose of injectable to a local holiday or event, and to write it down in the appointment card. Thus, as expected, it was found that the intervention area study participants were more likely to recall being told about a local holiday as a reminder for the appointment date. Associating the return appointment date with a local event or holiday seemed to be a more common practice among the intervention area HEWs, even without the availability of printed appointment cards. Recall of the side effects of injectable contraceptives was also higher among the intervention area participants, suggesting the possible effect of the client counseling job aid. Although L10K provided refresher training to HEWs in both study arms on the family planning client care job aid provided by the national program, this job aid was long and cumbersome for counseling clients. The behavioral design process made the original job aid more concise by increasing the salience of key messages and reducing the amount of information included. The new job aid was mounted on a desktop tower to be readily accessible during client counseling. These two design adaptations may have made the job aid more effective.

It is interesting to note that using the appointment card and the client counseling job aid were within the routine workflow of the HEWs, but the planning calendar required an alteration of the HEWs’ workflow. Thus, HEWs complied with the use of appointment cards and client counseling job aid, but not with the use of the planning calendar.

The study design cannot distinguish whether the appointment card or the counseling job aid was more effective. Ideally the tools could have been tested one by one, with a shorter duration of observation, to identify the effectiveness of each. The one-year observation period required to measure the outcome of interest for the study prevented the use of rapid tests one by one. This could be a blessing, because it is possible that the appointment card and counseling job aid had a synergistic effect. Nonetheless, a shorter testing duration is still essential to keep policy-makers interested, and to convert research findings into policy rapidly.

**Conclusion**

Alongside behavioral factors, injectable discontinuation can be influenced by a combination of health system factors (for example, injectable supply and availability of HEWs). The study’s long duration provided an opportunity for health system factors to mask the effects of the BE interventions.

Nonetheless, the HEWs in the intervention area consistently used two of the three BE tools, which overcame inconsistent health system performance in the health posts. The overall findings of this study suggest that BE is a promising approach for enhancing family planning programs in this low-resource, rural setting in Ethiopia.

**Consent**

For the FP client follow-up survey, verbal consent was sought and documented by the interviewers. If the respondent was under 18 years old, then consent was sought from her husband, parents or guardian. As it was expected that most of the respondents could not read or write, written consent was not sought. Only when consent was given by the respondent, the electronic data collection tool allowed to continue with the survey.

Data collection from health posts was considered a structured routine supportive supervisory visit. Data was collected mainly through observations and review of health post records and extraction of administrative data. However, verbal consent from the HEW or HEWs (when more than one HEW was present during health post visit) were obtained before continuing with data collection from the health post.
The consent forms for the FP client follow-up and health posts data collection were part of the data collection tool and can be accessed with the data repository.

Ethical approval was obtained from the ethical review the ethical review boards of the regional health bureaus of Amhara, Oromia, SNNP and Tigray regions and from the ethical review board of JSI Research & Training Institute, Inc. The ethical approvals can be accessed with the data repository.

This study was retrospectively registered with the International Standard Randomised Controlled Trials Number registry (ISRCTN17390653) on the 10 April 2019.

**Data availability**

**Underlying data**


This project contains the following underlying data:

- **BE health post fidelity.csv** (Health post survey data in comma-separated values format)
- **BE_client_interview.csv** (Study participants’ interview data in comma-separated values format)
- **hp_pop2018.csv** (Contains the population sizes of the health post catchment areas in comma-separated values format)
- **master.do** (Stata do- or syntax-file that implements the rest of the Stata do-files sequentially)
- **hp_InterventionFidelity.do** (Stata do-file that reproduces the analysis presented in Table 3 and creates the independent variables of Table 6 and also used for the multi-variate analysis presented in Table 7)
- **BE_client_interview1.do** (This and the next Stata do-file reproduces the analysis presented in Table 5, Table 6 and Table 7)
- **BE_client_interview2.do**

**Extended data**


This project contains the following extended data:

- **ClientInterviewQuestionnaire.pdf** (Questionnaire used to interview the study participants)
- **HealthPostQuestionnaire.pdf** (Questionnaire used to obtain data from the health posts of the study area)
- **Tables A1 and A2.docx** (Table A1 presents person-months of retrospective follow-up of the study participants by the independent variables considered for the discrete hazard model that estimated injectable discontinuation rate. It also includes the name of the variables constructed for analysis. Table A2 presents the independent variables used in each of the models presented in Table 7)

**Reporting guidelines**


Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

**Grant information**

This work was supported by the Bill and Melinda Gates Foundation [OPP1131042].

This work was also supported by the William and Flora Hewlett Foundation [2015-2551].

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

**Acknowledgement**

We thank the Health Extension Workers and their supervisors in the intervention and control areas for their contribution towards implementing the intervention. We are also thankful to the family planning clients who voluntarily participated in this study. The support from the Regional Health Bureau and L10K staff of Amhara, Oromia, SNNP and Tigray regions was critical to implement the study. We would like to acknowledge Teis Jorgensen and Saugato Datta, ideas42, for their contribution in conceptualizing and designing the study; Joan Robertson, JSI, for her contribution in acquiring funding and managing the grant; Esther Braud, JSI, for managing the grant; and Tahmid Chowdhury for visualization.

**References**


Dobility Inc: SurveyCTO. Cambridge, MA. 2018. Reference Source
StataCorp: Stata: Release 15, Statistical Software. College Station, TX: StataCorp LLC. 2017.
Open Peer Review

Current Peer Review Status: ? ?

Version 1

Reviewer Report 08 July 2019

https://doi.org/10.21956/gatesopenres.14094.r27399

© 2019 McConnell M. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Margaret McConnell
Department of Global Health and Population, Harvard T.H. Chan School of Public Health, Boston, MA, USA

I commend the researchers on an important study and well-written manuscript. I provide a few suggestions below. I believe these should be readily addressable:

1. In my mind it would be useful to more thoroughly discuss other relevant literature on reminders, job aids and the building blocks of the intervention more thoroughly (particularly those that take a similar behavioral approach). This paper by McConnell et al. (2018) is one example but placing the intervention within a broader context in both the introduction and discussion section would be valuable.

2. I am not clear on a few key elements of the design and analysis. First, which outcomes were registered as primary? Second, how does the analysis account for the clustering in the randomization? It would be useful to try different methods here including potentially randomization inference to obtain accurate standard errors.

3. It would be helpful to discuss potential endogeneity in the sample selection process. For example, if the interventions reduced discontinuation this would seem to suggest that the sample of individuals enrolled into the study would differ between arms as discontinuation was one of the criteria for eligibility. It would be useful to clarify and discuss this issue.

4. It appears that the rate of discontinuation was higher than expected. It would be helpful to discuss this more explicitly and to place in the context of the broader literature.

I also have some more minor points:

- In response to this sentence: “The possible bias due to loss to follow-up was minimized by the study design and by the application of regression methods” - I'm not clear on why regression methods would help with bias?
“This was not surprising” seems too editorial for a results section - would suggest editing this to just include findings.

References

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Global Health, behavioral economics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 12 Nov 2019
Ali Karim, Bill & Melinda Gates Foundation, Addis Ababa, Ethiopia

Point-by-point response to Margaret McConnell, Department of Global Health and Population, Harvard T.H. Chan School of Public Health, Boston, MA, USA

I commend the researchers on an important study and well-written manuscript. I provide a few suggestions below. I believe these should be readily addressable:
1. In my mind it would be useful to more thoroughly discuss other relevant literature on reminders, job aids and the building blocks of the intervention more thoroughly (particularly those that take a similar behavioral approach). This paper by McConnell et al. (2018) is one example but placing the intervention within a broader context in both the introduction and discussion section would be valuable.
Response: Agreed. Updated the introduction and discussion sections to place the current study within the broader context of the behavioral science literature. In the introduction, we clarified that behavioral barriers were likely at play in Ethiopia and discussed recent relevant behavioral economics literature that informed the building blocks of the intervention. In the discussion, we identified other behavioral science research to help explain the results of the study.

1. I am not clear on a few key elements of the design and analysis. First, which outcomes were registered as primary? Second, how does the analysis account for the clustering in the randomization? It would be useful to try different methods here including potentially randomization inference to obtain accurate standard errors.

Response: The primary outcome is injectable contraceptive discontinuation rate within 12 months of uptake. To make it explicit, we made edits throughout the manuscript and now refer to the other outcomes as secondary instead of intermediate outcomes.

The clustering of the randomization units (i.e., PHCUs or districts) is accounted for in the models we specified. Since there were only two randomization units within each region and one of them was randomly allocated to the intervention arm, the indicator variables for administrative regions and study arm accounted for PHCU-level clustering (see PHCURandomEffects.log of the extended data (Karim 2019)). Thus, based on the conclusion that the combination of administrative region and study arm indicator variables explained the PHCU-level unobserved variance, we estimated models considering only health posts as random factors. Nonetheless, we now use robust standard errors if the health post-level random-effects model were not significant. We update the analysis section and Table 7 accordingly.

1. It would be helpful to discuss potential endogeneity in the sample selection process. For example, if the interventions reduced discontinuation this would seem to suggest that the sample of individuals enrolled in the study would differ between arms as discontinuation was one of the criteria for eligibility. It would be useful to clarify and discuss this issue.

Response: The discussion on sample selection bias is now included in the discussion section. Specifically, we added the following: “Fourth, the design of participant recruitment and eligibility may have introduced sample selection bias. Specifically, participants were eligible for the study if they were new injectable users (defined as someone new to family planning, switching from one method to another, or who had taken a hiatus from family planning for six months or more). Because the intervention reduced discontinuation and encouraged dissatisfied users to switch methods, it is possible that the study participants who were included during the latter part of the enrollment period were different between the study arms. The intervention arm may have included more women who were switching to injectable contraceptives from a prior method. Tables 5 and 6 show that for both individual- and provider-level characteristics, there is broad balance between treatment and control participant. However, we are unable to completely eliminate the possibility of this bias.”

1. It appears that the rate of discontinuation was higher than expected. It would be helpful to discuss this more explicitly and to place in the context of the broader literature.

Response: In the discussion section, we now note that the discontinuation rate observed in our study was much higher than expected. We also provide a few insights on it. Specifically, we note: “Injectable discontinuation rates in both arms of the study were
much higher than initially expected. The difference in the representativeness and measurement method of the discontinuation rate between the expected (nationally representative data with a five-year recall period) and the observed (which was in eight districts with one-year recall period) may partly explain it. It could be that the respective Regional Health Bureaus purposely selected low performing districts for L10K to support the family planning program. Nonetheless, variations in the injectable discontinuation rate between geography is observed in Ethiopia. An independent study in 2014 in Agarfa district, Oromia, report significantly higher rates of injectable discontinuation (67%) than the national average (Bekele et al., 2015).

I also have some more minor points:

- In response to this sentence: "The possible bias due to loss to follow-up was minimized by the study design and by the application of regression methods" - I'm not clear on why regression methods would help with bias?

Response: Agree, we do not know if any of the bias due to loss to follow-up was accounted for by the regression models. We delete the sentence.

- "This was not surprising" seems too editorial for a results section - would suggest editing this to just include findings.

Response: Agree. Updated the results section of the abstract.

**Competing Interests:** No competing interests were disclosed.

Reviewer Report 02 July 2019

https://doi.org/10.21956/gatesopenres.14094.r27396

© 2019 Wesson J. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Jennifer Wesson

IntraHealth International, Chapel Hill, NC, USA

- The article is a combination of the description of how they used behavioral economics to design an intervention and how successful the intervention was. This is my main concern/confusion about the paper. Is it a paper highlighting BE, or a paper highlighting a successful intervention package? Sometimes it seems it is one way and sometimes it is the other way. For example, the discussion section largely describes the effect of the intervention and not the implications of using BE. Whereas, the conclusion largely seems to be about using BE approaches. If the paper is going to cover both subjects, then the discussion and conclusion should, as well.

- The discussion section is weak, in my opinion. It is largely just a repeat of what the results are. The discussion section is always the hardest to write. There doesn't seem to be a lot of original thinking in this paper's discussion. What does it mean that the intervention was statistically significant? Are
there more provider-behavior issues than client-behavior? How does the intervention tie back to the diagnosis?

- One question I have is whether clients agreed to be contacted by the study personnel. It sounds as if data collectors went through client registers and then hunted the clients down. This seems to be a breach of the clients’ privacy. I know the study received ethical approval, but I would think about this question going forward in other research studies.

- Finally, you could delete tables 5 and 6. Their import could be said in just a couple of sentences.

- All in all, a very interesting paper. I enjoyed reading it.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Family planning, epidemiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 12 Nov 2019


Point-by-point responses to Jennifer Wesson, IntraHealth International, Chapel Hill, NC, USA

- The article is a combination of the description of how they used behavioral economics to design an intervention and how successful the intervention was. This is my main concern/confusion about the paper. Is it a paper highlighting BE, or a paper highlighting a successful intervention package? Sometimes it seems it is one way and sometimes it is the other way. For example, the discussion section largely describes the effect of the
intervention and not the implications of using BE. Whereas, the conclusion largely seems to be about using BE approaches. If the paper is going to cover both subjects, then the discussion and conclusion should, as well.

Response: We have revised the abstract and updated the discussion and conclusion sections to reflect that the article is a combination of a) how we used behavioral economics to design intervention and b) how successful the intervention was.

- The discussion section is weak, in my opinion. It is largely just a repeat of what the results are. The discussion section is always the hardest to write. There doesn't seem to be a lot of original thinking in this paper's discussion. What does it mean that the intervention was statistically significant? Are there more provider-behavior issues than client-behavior? How does the intervention tie back to the diagnosis?

Response: The discussion section has been expanded to further explain the intervention results, the associated provider and client-side behaviors, the relevance of diagnosis findings, and areas for future exploration.

- One question I have is whether clients agreed to be contacted by the study personnel. It sounds as if data collectors went through client registers and then hunted the clients down. This seems to be a breach of the clients' privacy. I know the study received ethical approval, but I would think about this question going forward in other research studies.

Response: Written or verbal consent, depending on literacy, to participate in the study was sought from all family planning clients visiting the health post in both the study arms during the recruitment period (February 2016-November 2016). However, the consent records were unable to be retrieved from health posts. Thus, consent was sought again during the follow-up survey, which could have breached client privacy in a few cases. It is an important lesson for future research to take measures to maintain client consent records.

- Finally, you could delete tables 5 and 6. Their import could be said in just a couple of sentences

Response: We debated whether to keep the tables in the body of the paper or move them to the Appendix and ultimately decided to keep them in the main manuscript as it is likely that some of the paper’s audience would like to see an assessment of the balance between the two study arms to determine randomization validity.

- All in all, a very interesting paper. I enjoyed reading it.

Response: Thanks!

Competing Interests: No competing interests were disclosed.
Alison M. Buttenheim, University of Pennsylvania, USA

Thank you for the opportunity to comment on this manuscript, which reports results of a randomized trial of a set of innovative, behaviorally-informed strategies to improve injectable contraceptive method continuation rates in rural Ethiopia. The manuscript is particularly strong in its description of the intervention design process. Researchers working in similar or adjacent areas will be able to glean important lessons about how specifically to apply a behavioral design approach to selecting an appropriate behavioral target, identifying bottlenecks to that behavior, and how to design interventions that address these bottlenecks.

For the most part, the manuscript is very thorough and complete, and all data and code have been made available.

One area where more detail would be helpful is the "behavioral diagnosis Step 2" section on page 5. The process used to generate behavioral bottlenecks is somewhat glossed over - an Appendix with more process details or some worked examples would strengthen reproducibility. Table 1 is helpful in terms of identifying the design implications of each bottleneck, but it is less clear where the bottlenecks came from.

I have some statistical/analysis questions that could be addressed in the manuscript for greater clarity:

- **Outcome measurement discussion:** This important section is hard to follow. It is not clear whether outcome measurement takes into account the duration of contraceptive protection offered by the injection. In other words, does the trial measure receipt of injection, or duration of contraceptive coverage?
- **Sample size calculation vs. observed discontinuation rate:** The discussion section should comment on the fact that power calculations were done using 34% discontinuation rate as the baseline value, and the trial was powered to detect a 14% difference with 330 per arm. In the trial, a 53% discontinuation was observed in the control condition and 42% in the treatment group, with 408 and 335 participants. It's surprising with these observed values and sample size that you're powered to see a significant 11% difference? Could a posthoc power calculation confirm this?
- **Statistical analysis section, description of discrete time hazard model:** I would recommend a richer description of how the discrete-time hazard model results are converted into discontinuation rates to ease interpretation.
- **Discussion of Table 7 results:** Richer discussion is needed about the fact that the modal participant in the control condition received an appointment card and was told about a local holiday or event to anchor return revisit (60% for both measures). Given this, I think it's a stretch to attribute much of the drop in discontinuation to the appointment cards, even if redesigned and protocolized.
- **Discussion section, p.13:** The issue of false identities is important given how many clients were lost to follow-up (Table 4, third row) and the difference in erroneous record proportion in treatment vs. control. One could imagine either higher or lower rates of giving false identities in the treatment condition, depending on when in the visit women provide identifying information. Would a sensitivity analysis be possible to assess the degree of potential bias here? In other words, how much higher would discontinuation have to be among the intervention participants who could not be located to erase the observed effect? Or could you conduct an ITT analysis assuming that women lost-to-follow up discontinued? This would be helpful in assessing how robust the results are.

And a few more minor comments:
- Client counseling job aid: When this is first introduced on p.5, it should be noted that there was an existing job aid already in use.
- Trial implementation paragraph, p.5 second column: This paragraph needs editing for clarity and typos.
- Conclusion: All of the material currently in the conclusion section except the final sentence should be moved up into the general discussion section.
- Figure 5: I think there is a typo in the note - should be "(p<.05)".

Overall I think this is an excellent contribution to both the contraceptive coverage literature and the behavioral science literature. I plan to include the paper in my syllabus for a behavioral economics and health course.

**Competing Interests:** I have served as a paid consultant to ideas42 on several projects related to maternal-child health and contraceptive coverage in the past 3 years. I have collaborated with two authors (Guichon, Lorenzana) in the course of these consultations, and provided statistical input on a different trial to a third author (Barofsky). I was not directly involved in the study described in the present paper in any way.