The impact of a demand-side sanitation and hygiene promotion intervention on sustained behavior change and health in Amhara, Ethiopia: a cluster-randomized trial

Matthew C. Freeman
Maryann G. Delea
Jedidiah S. Snyder
Joshua V. Garn
Mulusew Belew
Bethany A. Caruso
Thomas F. Clasen
Gloria D. Scalar
Yihenew Tesfaye
Mulat Woreta
Kassahun Zewdu
Abebe Gebremariam

1 Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA; *matthew.freeman@emory.edu (MCF); mdelea@emory.edu (MGD); jedidiah.snyder@emory.edu (JSS); gloria.sclar@emory.edu (GDS); thomas.f.clasen@emory.edu (TFC)
2 School of Community Health Sciences, University of Nevada, Reno, NV, USA jgar@unr.edu (JVG)
3 Emory Ethiopia, Bahir Dar and Addis Ababa, Ethiopia; limulusewb@yahoo.co.uk (MB);
mulatworeta@gmail.com (MW); lekassahun4@gmail.com (KZ); abebegmariam@yahoo.com (AG);
4 Hubert Department of Global Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA; bcanuso@emory.edu (BAC)

5 Department of Social Anthropology, Bahir Dar University, Bahir Dar, Ethiopia; yihenewtesfaye@gmail.com (YT)

6 School of Nursing, Emory University; abebegmariam@yahoo.com (AG)

*Corresponding author: Matthew C. Freeman - matthew.freeman@emory.edu
ABSTRACT

Background: Behaviors related to water, sanitation, and hygiene (WASH) are key drivers of infectious disease transmission, and experiences of WASH are potential influencers of mental well-being. Important knowledge gaps exist related to the content and delivery of effective WASH programs and their associated health impacts, particularly within the contexts of government programs implemented at scale.

Methods: We developed and tested a demand-side intervention called Andilaye, which aimed to change behaviors related to sanitation, personal hygiene, and household environmental sanitation. This theory-informed intervention was delivered through the existing Ethiopian Health Extension Programme (HEP). It was a multilevel intervention with a catalyzing event at community level and behavior change activities at group and household level. We randomly selected and assigned 50 kebeles (sub-districts) from three woredas (districts), half to receive the Andilaye intervention, and half the standard of care sanitation and hygiene programming (i.e., community-led total sanitation and hygiene [CLTSH]). We collected data on WASH access, behavioral outcomes, and mental well-being.

Results: A total of 1,589 households were enrolled into the study at baseline; 1,472 households (94%) participated in an endline assessment two years after baseline, and approximately 14 months after the initiation of a multi-level intervention. The intervention did not improve construction of latrines (prevalence ratio [PR] 0.99; 95% CI: 0.82, 1.21) or handwashing stations with water (PR: 0.96; 95% CI: 0.72, 1.26), or the removal of animal feces from the compound (PR=1.10; 95% CI: 0.95, 1.28). Nor did it impact anxiety (PR 0.90; 95% CI: 0.72, 1.11), depression (PR: 0.83; 95% CI: 0.64, 1.07), emotional distress (PR: 0.86; 95% CI: 0.67, 1.09) or well-being (PR: 0.90; 95% CI: 0.74, 1.10) scores.

Discussion: We report limited impact of the intervention, as delivered, on changes in behavior and mental well-being. The effectiveness of the intervention was limited by poor intervention fidelity. While sanitation and hygiene improvements have been documented in Ethiopia, behavioral slippage, or
regression to unimproved practices, in communities previously declared open defecation free is widespread. Evidence from this trial may help address knowledge gaps related to challenges associated with scalable alternatives to CLTSH and inform sanitation and hygiene programming and policy in Ethiopia and beyond.

Trial registration: This trial was registered with clinicaltrials.gov (NCT03075436) on March 9, 2017.
INTRODUCTION

Inadequate water, sanitation, and hygiene (WASH) are key drivers of infectious disease transmission (Freeman et al., 2017; Wolf et al., 2018). Diarrhea accounts for an estimated 1.65 million deaths annually [5] and nearly 10% of all under-5 deaths in low-income settings [6]. Deficiencies in WASH are also a major contributor of neglected tropical diseases (NTDs) [7,8]. Over one billion people are at risk of soil-transmitted helminthiasis, which leads to nearly five million disability adjusted life years (DALYs), and schistosomiasis leads to two million DALYs [9,10]. Trachoma, the leading infectious cause of blindness [11], is precipitated by repeat infections with Chlamydia trachomatis bacteria, which are often perpetuated by poor hygiene [12]. These infections are environmentally mediated [13], and are largely attributed to inadequate WASH [14,15].

While WASH studies have primarily focused on infectious diseases or anthropometric measures of growth amongst young children, this narrow focus does not fully encapsulate the World Health Organization (WHO)'s definition of health as “a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity” [2]. A growing body of research has identified linkages between water and sanitation and mental health outcomes [2,16–18]. For example, extensive qualitative and quantitative research has demonstrated how water insecurity can influence mental health, particularly among women [19–22]. Research on sanitation and mental health is emergent, and predominantly qualitative [23–25,25]. A cross-sectional study in Odisha, India, found women's sanitation insecurity—their negative sanitation experiences and concerns—to be associated with stress, depression, distress, and impaired general well-being, even among those with access to a sanitation facility [26]. Further, a systematic review of sanitation and well-being found open defecation and use of sanitation facilities can negatively influence mental and social well-being for women and girls, especially when they experience or perceive a lack of privacy and safety [2]. As such, improvements in women's mental health, likely require more than physical access to sanitation facilities, but also gender-sensitive
modifications to facilities and shifts in gender norms to improve women’s experiences of sanitation [24,26]. To date, limited research has assessed the impact of water interventions on mental health outcomes [27], and few studies have assessed the impact of sanitation interventions on mental health outcomes [2].

Despite the urgent need to improve sanitation and hygiene — including the target of universal basic access to sanitation as part of Sustainable Development Goal target 6.2 — many large-scale sanitation interventions have shown poor uptake and sustainability [28], as well as mixed impacts on health [1]. Without sustained sanitation and hygiene behavior change, health gains are unlikely.

Community-led total sanitation (CLTS) has been heralded as a low-cost approach to improve community coverage of sanitation [29]. CLTS uses a demand-side approach—promoting the demand to execute improved sanitation behaviors, rather than supply-side provision of infrastructure—that involves engaging the community, typically via an initial “triggering” event, to become open defecation free (ODF) through community-activities and local champions. Rigorous evaluations of CLTS, like those of other sanitation interventions, have yielded mixed health effects [30–32]. To date, there is mixed evidence on the potential of CLTS to achieve and sustain changes to WASH coverage and access [33–36]. Engaging local champions in CLTS delivery may yield beneficial results. Program delivery through Health Extension Workers (HEW) and the engagement of teachers both led to substantial improvements in sanitation coverage and use [34], although less than when delivered by trained natural leaders in Ghana [33]; yet these gains were not well sustained [37].

There are several documented limitations of community-led total sanitation and hygiene (CLTSH), a variation of CLTS that incorporates hygiene-related interventions. HEWs charged with implementing CLTSH have many responsibilities, limited incentives and motivations, few tools, and little capacity to continually reinforce messages [38]. The use of negative affective motivators employed by CLTS(H) may
not be culturally appropriate or the most effective drivers of sanitation and hygiene behavior change [39], and may erode mental well-being. Together, the focus on negative affective motivators, poor facilitation of initial triggering, and a lack of follow-up, has left many communities with negative impressions of CLTSH initiatives [40].

In Ethiopia, which has been implemented widely through the Ethiopian Health Extension Programme (HEP), relies chiefly on negative affective motives (e.g., shame, disgust) to drive open defecation cessation. However, like prior evaluations of CLTS, evidence suggests that CLTSH is largely ineffective, with one out of six Ethiopian households continuing to practice open defecation after their respective villages were certified as ODF [41].

We designed a study to generate evidence to address knowledge gaps related to demand-side sanitation and hygiene programming and examine less studied, yet critical, inter-personal factors related to sustained behavioral adoption and downstream health impacts [42]. Specifically, we conducted a cluster-randomized trial (CRT) to test whether an intervention delivered at scale within the existing Ethiopian HEP would lead to sustained WASH behavior change and improved mental health. Leveraging feedback received from community members and key stakeholders, we designed a theoretically-informed [43–45] and evidence-based demand-side sanitation and hygiene intervention called Andilaye—Amharic for “togetherness/integration.” The intervention takes a positive, encouragement approach to behavior change by promoting incremental improvements in behavior and incorporating behavioral maintenance strategies to foster sustained behavior change.

METHODS

The study’s primary aim was to determine whether a demand-side sanitation and hygiene intervention (Andilaye) impacted WASH behavior change and mental health, specifically general well-being and
symptoms of anxiety, depression, and non-specific emotional distress. A protocol detailing the methods, intervention, and baseline results are published elsewhere [42].

Ethics and trial registration

Ethical approval for the Andilaye Trial was provided by Emory University (IRB00076141), the London School of Hygiene & Tropical Medicine (9595), and locally by the Amhara Regional Health Bureau (HRTT0135909). The trial was registered with clinicaltrials.gov (NCT03075436) on March 9, 2017.

Study design

This parallel CRT was conducted in West Gojjam and South Gondar Zones of the Amhara National Regional State, a region of Ethiopia in which WASH conditions are inadequate [46], slippage in sanitation coverage and improved sanitation behaviors has been documented, and several NTDs (e.g., soil-transmitted helminths, trachoma) are hyperendemic [47]. The specific behaviors targeted by the Andilaye intervention and their respective key outcome indicators are detailed in Table 1. We sought to investigate whether any changes in WASH behaviors targeted by the Andilaye intervention were sustained, and we tracked intervention fidelity through a process evaluation.
Table 1. Key outcome indicators for WASH behavioral themes and constituent practices of interest of the Andilaye intervention at endline

<table>
<thead>
<tr>
<th>Indicators</th>
<th>Intervention</th>
<th>Control</th>
<th>PR (95% CI)</th>
<th>PD (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sanitation (S)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>S1: Construct a long-lasting latrine that is comfortable and hygienic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Households with access to at least one household latrine</td>
<td>743</td>
<td>61.2</td>
<td>729</td>
<td>62.0</td>
</tr>
<tr>
<td>- Households with access to an improved household latrine</td>
<td>741</td>
<td>34.6</td>
<td>726</td>
<td>30.6</td>
</tr>
<tr>
<td>- Households with access to a fully constructed household latrine</td>
<td>742</td>
<td>33.0</td>
<td>729</td>
<td>28.7</td>
</tr>
<tr>
<td><strong>S2: Repair your latrine whenever it is damaged</strong></td>
<td>455</td>
<td>70.1</td>
<td>451</td>
<td>80.5</td>
</tr>
<tr>
<td>- Facility observed to require obvious repair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>S3: Upgrade your latrine so it becomes more long lasting, comfortable, and hygienic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Households with added or improved anything on the latrine since its original construction</td>
<td>453</td>
<td>17.2</td>
<td>446</td>
<td>15.7</td>
</tr>
<tr>
<td>- Households with latrine with smooth and cleanable slab/floor</td>
<td>743</td>
<td>16.3</td>
<td>728</td>
<td>13.3</td>
</tr>
<tr>
<td>- Presence of drop hole cover in the latrine</td>
<td>455</td>
<td>18.2</td>
<td>451</td>
<td>10.0</td>
</tr>
<tr>
<td><strong>S4: Close your pit when it becomes full and reconstruct a new latrine</strong></td>
<td>454</td>
<td>11.7</td>
<td>451</td>
<td>12.6</td>
</tr>
<tr>
<td>- Is the pit that is in use full or close to being full</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>S5: All household members use a latrine every time they defecate</strong></td>
<td>743</td>
<td>53.2</td>
<td>729</td>
<td>54.1</td>
</tr>
<tr>
<td>- Respondent always exclusively used a latrine for defecation during last 7 days</td>
<td>529</td>
<td>36.5</td>
<td>473</td>
<td>33.0</td>
</tr>
<tr>
<td>- Head of household always exclusively used a latrine for defecation during last 7 days</td>
<td>1447</td>
<td>42.6</td>
<td>1385</td>
<td>35.0</td>
</tr>
<tr>
<td><strong>S6: Immediately dispose of children’s feces into the latrine</strong></td>
<td>401</td>
<td>36.7</td>
<td>376</td>
<td>41.2</td>
</tr>
<tr>
<td>- Child feces were safely disposed of during the last 2 days</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Personal hygiene (PH)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PH1: All household members wash their hands with water and soap or soap substitute AFTER handling animal and human feces, even children’s feces</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Household hand or facewashing station(s)</td>
<td>743</td>
<td>98.3</td>
<td>729</td>
<td>97.7</td>
</tr>
<tr>
<td>- The last time the respondent defecated, s/he cleaned hands with water and soap, substitute</td>
<td>738</td>
<td>51.9</td>
<td>725</td>
<td>46.1</td>
</tr>
<tr>
<td>- The last time the index child defecated, s/he cleaned hands with water and soap, substitute</td>
<td>713</td>
<td>43.9</td>
<td>687</td>
<td>39.6</td>
</tr>
<tr>
<td><strong>PH2: All household members wash their hands with water and soap or soap substitute BEFORE handling food</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- The last time the respondent prepared food, s/he cleaned hands with water and soap, substitute before beginning food preparations</td>
<td>700</td>
<td>53.6</td>
<td>703</td>
<td>48.5</td>
</tr>
<tr>
<td><strong>PH3: All household members wash their faces with water whenever they are dirty and use soap when it is available</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
- Ocular discharge present among children aged 1-9 years
  822  26.9  874  30.4  0.88 (0.68, 1.15)  -0.037 (-0.114, 0.040)
- Wet nasal discharge present among children aged 1-9 years
  822  37.0  874  39.4  0.94 (0.78, 1.13)  -0.024 (-0.094, 0.045)
- Dry nasal discharge present among children aged 1-9 years
  822  42.7  874  45.2  0.97 (0.81, 1.16)  -0.016 (-0.095, 0.063)
- Dirt/dust/other debris present among children aged 1-9 years
  822  50.5  874  49.5  1.03 (0.89, 1.20)  0.016 (-0.058, 0.090)

Household Environmental Sanitation (HES)

**HES1:** Keep all animals separated from the house
- Observed animal feces present in the compound
  743  82.2  729  82.4  1.01 (0.92, 1.11)  0.0071 (-0.07, 0.08)

**HES2:** Keep the household compound clean by disposing of all animal feces and other waste on a DAILY basis
- Animal feces/waste not left out in open in compound
  743  56.4  729  51.2  1.10 (0.95, 1.28)  0.052 (-0.029, 0.132)
- Solid waste was not observed to have been left out in the open
  743  34.6  729  27.6  1.26 (0.93, 1.69)  0.0705 (-0.02, 0.17)

**Notes.**

- We used log-linear binomial regression models to compare the prevalence of the outcomes between the intervention and control arms. Models accounted the stratified design by including woreda indicator variables [48], and accounted for clustering within kebeles by using generalized estimating equations with robust standard errors.
- Prevalence differences (PD) were calculated using post-estimation commands to estimate the average marginal effects.
- "Improved" was defined based on the WHO/UNICEF Joint Monitoring Programme (JMP) for Water Supply and Sanitation definition.
Fifty kebeles—the smallest government administrative unit in Ethiopia—in Amhara Region were selected and randomly assigned to either the Andilaye intervention or counterfactual (i.e., control comparison group). Kebele clusters, the randomization unit, were rural or peri-urban sub-districts from three districts (woredas)—Bahir Dar Zuria, Fogera, and Farta. These areas were targeted because they represented different topographical conditions present in Amhara. Following baseline data collection, we randomly selected and assigned clusters using stratified sampling by district. Randomization was done using a computer-based random number generator. To secure balance across three key potential confounders (i.e., latrine coverage, washing station with soap coverage, and head of household education), we established a priori that the intervention and control mean values for these three variables, using baseline survey data, should be within two standard deviations of the overall mean. As such, the randomization process was repeated twice using replacement rerandomization [49] to achieve balance according to that a priori criterion. Figure 2 provides further details in a flow diagram.

“Andilaye” Intervention

The Andilaye intervention motto was “Together we can be a strong, caring, healthy community”. Intervention activities offered aspirational messages that emphasized the need for collective action to make positive change in the community and used verbal persuasion to enhance collective efficacy perceptions [50]. The Andilaye intervention focused on three WASH-related behavioral themes, informed by formative research: (1) sanitation, (2) personal hygiene, and (3) household environmental sanitation. Within these themes were 11 constituent practices targeted by the intervention (Table 1); these practices were identified through formative research as ones that could be targeted using demand-side approaches, and were seen as achievable, per stakeholder feedback [42]. Intervention activities and behavior change tools were informed by our formative research and specifically designed to incorporate techniques that addressed behavioral factors such as action
knowledge, perceived personal and household barriers to behavioral adoption, identification and planning, and behavioral control perceptions amongst others [42]. Activities occurred at four levels—district, community, group, and household (Supplemental Figure 1 and Supplemental Table 1). Key activities included community mobilization and commitment events, community conversations, and household counseling visits with caregivers—all of which were guided by behavior change tools (e.g., community commitment banner, community conversations flipbook, and household counseling flipbook and goal cards) with illustrations produced by an artist based in Ethiopia.

The Andilaye intervention commenced with district-level capacity building activities, such as action planning and training of trainers who would facilitate intervention activities. Further, district-level refresher trainings and adaptive management activities were conducted to reinforce previously acquired knowledge and skills, address trainer/facilitator turnover, and review successes and address challenges faced in implementing group and household level activities. Community-level activities included the ‘Whole System in the Room’ [51], community mobilization and commitment events, and cross-fertilization visits. During household-level counseling visits, trained Women’s Development Army Leaders (WDALs) provided personalized counseling to caregivers to equip them with the knowledge, skills, and motivation necessary to adopt and maintain improved WASH practices. Structured community conversations, implemented by trained community facilitators, provided further opportunity for group-level counseling and support.

Figure 1. The Andilaye Trial consists of three major phases: (1) formative research and intervention design, (2) intervention implementation and process evaluation, and (3) impact evaluation. Kebele and household enrollment took place during baseline data collection (March to April 2017). Implementation of Andilaye intervention activities began in September 2017 and continued through midline data collection (March to April 2018), quarterly monitoring (June to July and November to December 2018), and endline evaluation (March to May 2019). See Supplemental Table 2 for specific dates of the delivery of intervention activities. Midline data reflected at least 2 months since the start of household-level behavior change activities and 3 weeks since the completion of a catalyzing community-level
mobilization and commitment event. Our endline data reflected the implementation of 14-15 months of household-level behavior change activities and 6-7 months of group-level behavior change activities (as dictated by trained activity facilitators) and 13-14 months since the community mobilization and commitment events.

The Andilaye intervention was delivered through Ethiopia's HEP, via trained government-salaried Woreda (district) Health Office officials, HEWs, and volunteer WDALs (Figure 1; see Supplemental Table 2 for an alignment of relevant roles and responsibilities of the HEP and Andilaye Trial). Implementation of the Andilaye intervention was overseen by an Ethiopian-based study team. The Health Extension Services Package, and its accompanying CLTSH module delivered via the HEP, were being scaled throughout Ethiopia [52] and reflected the existing government-supported demand-side sanitation and hygiene approach. No attempt was made to modify the roll out of this standard care of WASH programming in any of our study clusters. Results from our equivalence analyses indicated balance in the number of previously CLTSH-triggered and ODF certified kebeles, between study arms (Figure 2).

Baseline statistics, along with the fact that 39 of 50 kebele clusters randomly selected for inclusion in the Andilaye Trial had been triggered with CLTSH, and certified ODF, provided strong evidence that behavioral slippage was, indeed, an issue that needed to be addressed in Amhara and perhaps elsewhere in Ethiopia.

[Figure 2. Study flow diagram]

Outcomes of interest

Survey instruments administered for our impact evaluation collected data on key outcomes through self-reports from respondents and other household members. Primary outcomes of interest included mental health and three targeted WASH behavioral themes (1) sanitation, (2) personal hygiene, and (3)
household environmental sanitation behaviors, consisting of 11 constituent practices (Table 1). To
measures WASH outcomes, we pulled from standard WASH indicators, and leveraged formative
research data to contextually adapt survey prompts and answer choices (shown in Supplemental Table 4
of Delea et al., 2019). Sustainability of WASH-related behaviors was measured through the proportion of
individuals and households consistently practicing target behaviors at midline and endline.

For mental health, we assessed subjective well-being using the validated WHO’s Well-Being Index
(WHO-5) [53] and symptoms of anxiety, depression, and non-specific emotional distress using the
Hopkins Symptom Checklist (HSCL) [54]. The WHO-5 asks the respondent to indicate how frequently
they relate to each of five statements in the previous two weeks using a five-point Likert scale. Higher
scores are better (range: 0-25) with scores below 13 indicating poor well-being. The HSCL is a non-
diagnostic tool that includes 25 items to assess symptoms of anxiety (items 1-10), depression (items 11-
25) and overall emotional distress (all 25 items). We omitted two items from the depression set: an item
on sexual desire, which was deemed inappropriate for unmarried women, and an item on suicide
ideation, because we were unable to provide clinical recourse if needed. Participants indicated how
much symptoms bothered them in the previous week (‘not at all’ [1] to ‘Extremely’ [4]). The final score
for each state is a mean of responses for each of the relevant items (range from 1 to 4). Scores of 1.75 or
higher indicate that the condition could be present while lower scores are an indication of lower anxiety,
depression, or distress.

Secondary outcomes included diarrhea period prevalence, measured through caregiver report of the
index child (i.e., youngest child in the household aged one to nine years at baseline), sanitation
insecurity, and water insecurity. For sanitation insecurity, we asked respondents to indicate how often
(never, sometimes, often, always) they felt one of seven different forms of sanitation insecurity (i.e., 7
factors). Scores were means of all items in the factor. A higher score represents higher sanitation
insecurity. The factors were predesignated, and based on a validation that was done in another study [24]. Water insecurity was measured through the Household Water Insecurity Experiences (HWISE) scale [55]. HWISE includes 12 items with four response categories (never, rarely, sometimes, often/always). The score is the sum of responses, ranging from 0–36. A higher score indicates greater household water insecurity, and we considered households to be water insecure with a score of 12 or more, as described elsewhere [55]. Measures of collective efficacy [50] and intermediate behavioral antecedents were also collected, but are not reported here.

Process evaluation

We conducted a process evaluation alongside our impact evaluation to describe and analyze key aspects of the Andilaye’s implementation and provide insights and understanding of program impacts. We defined fidelity as the degree to which the intervention or program was delivered as intended [56]. Quantitative process data on dose delivered, participation, and dose received were collected through the direct observation of all district and community-level activities—these activities being facilitated or co-facilitated by the Ethiopian-based study team. Questions were incorporated into our survey instruments administered during quarterly monitoring and endline data collection to capture exposure to key Andilaye behavior change activities by respondents from study-enrolled households in intervention kebeles. This included self-reported awareness of and attendance at the community mobilization and commitment event and community conversations, and the number of household counseling visits received from WDALs. Per protocol, all community members were targeted to attend the community mobilization and commitment event; routine (1-2 per month) community conversations primarily focused on influential community members (e.g., male heads of households, religious leaders, mother-in-laws) targeted in the ‘Whole System in the Room’; and caregivers were to receive monthly
household counseling visits (each visit lasting approximately 30 minutes) following the typical structure for the WDAL and in accordance the HEP.

Sample size and power
A detailed description of sample size considerations is published in the study protocol [42]. Briefly, we powered this study on mental well-being outcomes, as measure by the HSCL [54], utilizing data from Ethiopia and East Africa suggesting that approximately 20-35% of rural women experience elevated symptoms of common mental disorders such as anxiety and depression [57,58]. Our sample size determination indicated we should recruit and enroll 25 households from each of 50 kebele study clusters, with 25 kebeles per study arm, targeting one index child—the youngest child aged one to nine years at baseline—per household. We increased our final sample size to accommodate for 20% of households being lost to follow-up. Our target sample, therefore, included 30 households in each kebele study cluster, or 1,500 households in total (i.e., 750 per study arm).

Data collection
Data were collected via structured household interviews and observations by trained enumerators during rounds of data collection. Surveys were collected using mobile phones equipped with the freely available Open Data Kit (http://opendatakit.org/). We interviewed the same households at each round of data collection. Figure 1 and Supplemental Table 2 provide a summary of the timeline of intervention implementation in relation to points of data collection for our impact evaluation.

Analytical methods
We followed a pre-analysis plan developed following baseline data collection [42]. The primary analysis method was an “intention-to-treat” analysis, which compares the intervention arm to the
counterfactual arm without regard to intervention fidelity or compliance. The majority of our primary and secondary outcomes were binary variables, and for these we used log-linear binomial regression models and report the prevalence ratio (PR). For these binary outcomes we also present prevalence differences (PD), which were calculated using post-estimation commands to estimate the average marginal effects. For continuous outcomes, such as the WHO-5 and HSCL scores, we used linear regression models. All models included an intervention variable as a fixed effect, account for the stratified design through the inclusion of the woreda indicator variable [59], and incorporated generalized estimating equations with robust standard errors to account for the clustering of observations within kebeles. For each of our primary outcomes of interest, we assessed if there was interaction across various sub-groups, including exposure to previous CLTSH triggering and sex of the index child. We also assessed if water insecurity modified the effectiveness on hygiene behaviors. For all of these analyses, we included interaction terms to test if effect modification was present (i.e., the interaction term had a p-value <0.05).

To assess whether any improvements in WASH behaviors were sustained between the midterm and the follow-up period, we compared the prevalence of key targeted sanitation, hygiene, household environmental sanitation indicators between the demand-side intervention arm and the counterfactual arm group using the baseline, midline, and endline data (Figure 1). Based on the endline data, at the time of data collection household-level activities had taken place for the past 14-15 months, group-level activities for the past 6-7 months, and community-level mobilization and commitment events were completed 13-14 months prior. Elements of the intervention were still ongoing during the time of our endline data collection, and we did not collect further data after the endline visit.

RESULTS:
Survey results

Our baseline assessment showed balance in terms of our primary outcomes of interest and demographic variables [42]. Our endline results reflect complete data from 1,472 (93%) of 1,589 households enrolled in the study at baseline, and was well within our sample size requirement of 1,250 households. Of the 793 enrolled intervention households and 796 enrolled counterfactual households, retention was similar in both arms, at 94% and 92%, respectively (Figure 2). A large majority (90%) of the respondents were female, by design. Of these 1,472 respondents, 85% were the mother of the index child.

Process evaluation

Reports from Woreda Health Offices collected at endline indicated that none of the 50 kebele study clusters (intervention or counterfactual) received additional CLTSH triggering or re-triggering during the course of the Andilaye Trial. The process evaluation survey was conducted at 703 randomly selected study-enrolled households in the months following the event (i.e., quarterly monitoring visits).

For the Andilaye intervention, the fidelity of action planning workshops and trainings at the district and community-levels were high (See Supplemental Table 3 for a summary of process data on dose delivered, participation, and dose received). In each district (woredas), we conducted a sensitization and action planning workshop, trainings of trainers and intervention activity facilitators (including one round of refresher trainings) and an adaptive management workshop. All intervention kebeles (n=25) had delivery of the ‘Whole System in the Room’ and action planning activity, community mobilization and commitment event, and skill-based training of WDALs (including two rounds of review meetings and refresher trainings).
Respondent-reported exposure of key behavior change activities varied, but was generally suboptimal (Figure 3). Overall, only 18% of respondents reported attending the community mobilization and commitment event, and 22% reported being aware of the activity, according to data collected from intervention household respondents via process evaluation surveys administered months following quarterly monitoring visits. WDALs and their supervisors (i.e., HEWs) were trained to facilitate monthly counseling visits with households in their catchment area. However, among intervention household respondents of process evaluation surveys administered at endline (n=665), only 59% reported receiving a counseling visit and 43% reported receiving at least one follow-up visit. No intervention kebeles had WDALs conducting counseling visits monthly. Among household respondents reporting at least one household visit (n=391), the average number of visits reported was 2-3 during the 14-15-month implementation period (Supplemental Table 3). Among respondents of the group-level process evaluation surveys administered at endline (n=707), 28% reported attending at least one community conversation, and 46% reported being aware of the activity.

[Figure 3. Respondent-reported exposure of key behavior change activities of the Andilaye intervention. Respondents from study-enrolled households in intervention kebeles (n=793) were surveyed on their awareness and attendance in the community-level mobilization and commitment event during quarterly monitoring visits and awareness and attendance of group-level community conversations and frequency household-level counseling visits received by endline visits. A total of 703 (89%) and 707 (89%) surveys with responses to process evaluation prompts were completed from quarterly monitoring and endline, respectively. Reported frequency of counseling visits reflected a total of 665 survey responses as these questions were not relevant for study-enrolled households that were residents of caregivers who were trained as Women’s Development Army Leaders (WDALs) responsible for conducting the Andilaye counselling visits (n=42).]

Impacts on sanitation, personal hygiene, and household environmental sanitation

The intervention did not increase latrine access. At endline, 62% of both intervention and control households had at least one latrine (prevalence ratio [PR] 0.99; 95% CI: 0.82, 1.21) (Table 1). There was no difference in the prevalence of improved latrines (PR 1.13; 95% CI: 0.81, 1.59) or in fully constructed
latrines (PR 1.15; 95% CI: 0.86, 1.54). Although there were improvements in many latrine characteristics in the intervention arm compared to the counterfactual arm (e.g., presence of water available or cleansing agent near or inside the latrine for handwashing, and water available for flushing or self-cleansing), the conditions (e.g., presence of feces on floor) of latrines in the intervention arm were often poor (Supplemental Table 4).

The intervention did not impact defecation practices. Overall, 40% of respondents reported practicing open defecation during the previous two days; only 46% of respondents had defecated in any latrine during the previous two days (Supplemental Table 4). All measures of latrine utilization and non-utilization were similar across intervention and counterfactual arms. This includes indicators of respondent open defecation (PR: 1.05; 95% CI: 0.76, 1.45), safe disposal of child feces (PR: 0.96; 95% CI: 0.69, 1.32), and number of people from another household who used a latrine during last seven days (difference: -0.40; 95% CI: -0.85, 0.05).

The intervention did not impact personal hygiene behaviors. The prevalence of washing stations with water (PR: 0.96; 95% CI: 0.72, 1.26) was similar between the intervention and counterfactual arms. Presence of hand or face washing stations were observed in 98% of households (Table 1), although water and soap were observed in only 20% of handwashing stations and 2% of facewashing stations (Supplemental Table 4). The prevalence of stations with soap was higher in the intervention arm, although only 3% of households in this arm had a washing station with soap present. Among all children aged 1-9 years at endline, observations of facial cleanliness indicated 29% had ocular discharge, 38% had wet nasal discharge, 44% had dry nasal discharge, and 50% had dust, dirt, or debris on their faces (Table 1). There were no meaningful differences between the study arms for any of these facial cleanliness measures.
We found no evidence that the Andilaye intervention impacted household environmental sanitation. Across both arms, the majority of respondents and heads of household had animal herding responsibilities (88% overall), and animal feces were present in the compound in 82% household compounds (Supplemental Table 4). A similar proportion of households in intervention and control kept animals separate from the house (PR=1.01, 95% CI 0.91, 1.11). About half of households did not leave animal feces/waste in the open (Table 1); this was similar between the intervention and counterfactual arms (PR=1.10; 95% CI: 0.95, 1.28).

**Sustained changes of key indicators**

There was little difference in the sustainability of key targeted indicators — assessed by changes between midline and endline — on sanitation access and practices, personal hygiene access and practices, and household environmental sanitation over the course of follow up (Figure 4). At midline, most variables continued to show little difference between the intervention and counterfactual arms, although the prevalence of drop hole covers in latrines and the prevalence of appropriate hygiene behaviors were more common in the intervention arm. At the endline visit, the prevalence of drop hole cover was largely sustained in the intervention arm, while the prevalence of drop hole covers decreased in the control arm (PR=1.77; 95% CI: 1.19, 2.63). All other variables at endline had similar prevalence levels when comparing the two arms. While the prevalence of hand hygiene behaviors was maintained at levels similar to the midline visit, increases in hand hygiene behaviors in the counterfactual arm narrowed the difference between the intervention and control arms at endline. Over the two follow-up surveys, there was an increase in the prevalence of household hand or facewashing stations that appeared among study arms.
**Figure 4.** The prevalence of key sanitation, hygiene, and environmental sanitation indicators over time.

**Impacts on mental health**

There was no difference between study arms in the scores for anxiety, depression, emotional distress or general well-being (Table 2). There was also no difference between the intervention and counterfactual arms in the prevalence of each mental health condition: anxiety (PR = 0.90; 95% CI: 0.72, 1.11), depression (PR: 0.83; 95% CI: 0.64, 1.07), emotional distress (PR: 0.86; 95% CI: 0.67, 1.09) and poor well-being (PR: 0.90; 95% CI: 0.74, 1.10) (Table 2).

**Table 2.** Mental well-being outcomes at endline

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Scores</th>
<th>Prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cronbach Alpha</td>
<td>Intervention</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Anxiety score b</td>
<td>0.89</td>
<td>742 1.46 (0.61)</td>
</tr>
<tr>
<td>Depression score b</td>
<td>0.87</td>
<td>742 1.35 (0.48)</td>
</tr>
<tr>
<td>Emotional distress score b</td>
<td>0.93</td>
<td>741 1.29 (0.46)</td>
</tr>
<tr>
<td>Well-being score c</td>
<td>0.97</td>
<td>749 17.6 (6.8)</td>
</tr>
<tr>
<td>High Anxiety f</td>
<td>-</td>
<td>742 22.2</td>
</tr>
<tr>
<td>High Depression f</td>
<td>-</td>
<td>742 14.0</td>
</tr>
<tr>
<td>High Emotional distress f</td>
<td>-</td>
<td>741 14.0</td>
</tr>
<tr>
<td>Poor well-being g</td>
<td>-</td>
<td>749 25.2</td>
</tr>
</tbody>
</table>

**Notes.** We used linear regression models to estimate the difference in the outcomes comparing the intervention and control arms. Models accounted for the stratified design by including woreda indicator variables (Kahan and Morris, 2012) and accounted for clustering within kebeles by using generalized estimating equations with robust standard errors. We asked respondents to indicate how much the symptoms bothered them in the previous week with four potential response options (not at all (1) to extremely (4)). The first ten symptoms assess anxiety (i.e., ‘suddenly scared for no reason’, ‘nervousness or shakiness inside’), the next 13 assess depression (i.e. ‘feeling low in energy’, ‘feeling hopeless about the future’), and the 23 collectively assess non-specific emotional distress. For each outcome, the score is the sum of the responses divided by the number of items. We asked respondents about well-being, and responses ranged from ‘(0) At no time’ to (5) All of the time’. Scores were summed, and range from 0-25; the higher the score, the better the well-being. We used similar log-linear binomial regression models to compare the prevalence of the outcomes between the intervention and control arms. Prevalence differences (PD) were calculated using post-estimation commands to estimate the average marginal effects. Each of the above scores was dichotomized, with scores greater than 1.75 indicating a positive status for any of the three outcomes. The above score was dichotomized with scores below 13 indicating poor well-being.
Secondary health outcomes

Reported diarrhea

Diarrhea prevalence during the last seven days among index children was similar in the intervention (7%) and counterfactual (6%) arms (PR: 1.20; 95% CI: 0.74, 1.93; Table 3). Among index children, there were also similarities comparing study arms in diarrhea prevalence when assessing episodes over the last two days (PR: 1.25; 95% CI: 0.71, 2.22).
Table 3. Secondary health outcomes at endline

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Cronbach Alpha</th>
<th>Intervention</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diarrhea</td>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>During the last 2 days, index child had three or more loose stools per day</td>
<td>-</td>
<td>730</td>
<td>5.9</td>
</tr>
<tr>
<td>During the last 7 days, index child had three or more loose stools per day</td>
<td>-</td>
<td>731</td>
<td>7.1</td>
</tr>
</tbody>
</table>

Water and sanitation insecurity scores

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Intervention</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water-HWISE Scale d</td>
<td>N</td>
<td>mean (SE)</td>
</tr>
<tr>
<td>Sanitation-Potential harms d</td>
<td>0.85</td>
<td>365</td>
</tr>
<tr>
<td>Sanitation-Social expectations resultant repercussions d</td>
<td>0.79</td>
<td>366</td>
</tr>
<tr>
<td>Sanitation-Physical exertion or strain d</td>
<td>0.57</td>
<td>366</td>
</tr>
<tr>
<td>Sanitation-Night concerns d</td>
<td>0.56</td>
<td>366</td>
</tr>
<tr>
<td>Sanitation-Social support d</td>
<td>0.88</td>
<td>366</td>
</tr>
<tr>
<td>Sanitation-Physical agility d</td>
<td>0.56</td>
<td>366</td>
</tr>
<tr>
<td>Sanitation-Defecation place d</td>
<td>0.81</td>
<td>366</td>
</tr>
</tbody>
</table>

Water insecurity prevalence

<table>
<thead>
<tr>
<th>Water insecure (HWise score 12 or more) d</th>
<th>N</th>
<th>%</th>
<th>N</th>
<th>%</th>
<th>PR (95% CI)</th>
<th>PD (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water insecure</td>
<td>-</td>
<td>565</td>
<td>5.7</td>
<td>388</td>
<td>8.8</td>
<td>0.50 (0.21, 1.2)</td>
</tr>
</tbody>
</table>

Notes. a We used log-linear binomial regression models to compare the prevalence of the outcomes between the intervention and control arms. Models accounted for the stratified design by including woreda indicator variables (Kahan and Morris, 2012) and accounted for clustering within kebeles by using generalized estimating equations with robust standard errors. b Prevalence differences (PD) were calculated using post-estimation commands to estimate the average marginal effects. c We used similar linear regression models to estimate the difference in the outcomes comparing the intervention and control arms. d We asked respondents to indicate how often they felt some form of sanitation insecurity (never, sometimes, often, always). These items were then summed with all other items in that factor and divided by the numbers of items to create a score. The factors were predesignated, and based on a validation that was done in another study (Caruso et al., 2017a). A higher score represents higher sanitation insecurity. e We used similar linear regression models to estimate difference comparing the outcomes between the intervention and control arms. This used a 12-item scale with four response categories (never, rarely, sometimes, often, always), and a total summed score of those response categories ranging from 0–36. A higher score indicates greater household water insecurity. We considered water insecure as a score of 12 or more, as described elsewhere scale [55].
Water and Sanitation insecurity

The intervention did not reduce water insecurity prevalence between intervention (5.7%) and counterfactual (8.8%) arms (PR: 0.50 (95% CI: .21, 1.2); Table 3). At endline, sanitation insecurity scores related to social support were statistically lower (i.e., better) in the intervention arm than in the control arm (score difference: -0.10, 95% CI: -0.16, -0.43), indicating a reduced frequency of experiencing the circumstances in the social support domain (e.g., trouble finding support to watch dependents during urination, worry about dependents when going to defecate, had to leave dependents alone to urinate, etc.). Other sanitation insecurity measures were similar between arms.

Interaction and effect modification

There was no interaction of the intervention by previous CLTSH triggering for any of the primary outcome variables of interest. We did not detect effect measure modification by sex for any of the four mental health outcomes. Similarly, we did not detect interaction by child’s sex for any of these outcomes. We also did not detect interaction between the intervention and water insecurity on any of the primary handwashing or face washing variables.

DISCUSSION:

The Andilaye intervention generally did not improve WASH conditions or outcomes. Without sustained changes to these WASH conditions and behaviors, changes in well-being were not likely, and indeed, were not detected. Improving sanitation and hygiene behaviors in rural communities remains a considerable challenge, especially in regions with poor water access and high levels of WASH-related NTD endemicity. Most studies designed to change sanitation and hygiene behavior are efficacy studies [28] – meant to assess changes under controlled conditions; ours was an effectiveness study, designed to measure changes in a real-world context. We believe that poor fidelity of intervention delivery played
a considerable role in uptake of the intervention, pointing to challenges in delivering demand-side sanitation and hygiene interventions at scale and through existing community-based models.

We did not find statistical differences between study arms at endline and few promising trends in intervention communities for some of the targeted behaviors. Changes were much lower than with approaches found by Crocker et al. elsewhere in Ethiopia [34,37]. Similarly, Apanga et al. found that implementation of the Rural Sustainable Sanitation and Hygiene for All (SSH4A) approach that employs a multidimensional intervention led to a large increase of 77 percentage points in sanitation coverage in Ethiopia and also coverage gains in many other countries under study [60], although slippage did occur after conclusion of program activities, whereas many other countries sustained their previous sanitation coverage gains [61].

Our Andilaye intervention did not impact validated mental health measures. Few studies have measured the impact of a sanitation intervention on mental health outcomes despite calls for broader investigations of sanitation-related health impacts [2,62,63]. We assessed if changes to sanitation access and sanitation insecurity — changes that we anticipated would be generated by this intervention — would lead to improved mental health states, including improved well-being and reduction in symptoms associated with anxiety, depression, distress and general wellbeing. In rural India, women's experiences of sanitation, as measured by a validated sanitation insecurity measure, were associated with well-being, anxiety, depression, and distress, even when women had access to a facility [26]. Similarly, in urban Mozambique, latrine location and neighborhood violence were important determinants of safety perceptions and corresponding psychosocial stress [64]. These findings highlight the need for interventions to consider the experience of sanitation beyond access to a facility alone. The intention-to-treat analysis did not detect changes to mental well-being scores or to sanitation insecurity scores, which was perhaps limited by our evaluation period (Figure 1). We believe further studies are warranted
to test the hypothesis that improved sanitation would impact mental well-being, as our intervention did not change sanitation behaviors, quality, or access.

The purpose of this study was to develop and test an intervention that could be scaled within the existing Ethiopian HEP. The intervention was designed to be incorporated into prevailing programs (e.g., HEP) to demonstrate potential for scale-up, and did not succeed in this regard. For example, despite strong attendance of trainings, action planning workshops, provision of supportive supervision and on-the-job-training tools, WDALs and HEWs did not conduct the necessary household visits. They reported that they did not receive supportive supervision from relevant government officials in accordance with their action plans (Supplemental Table 3). While supportive supervision considerations were acknowledged and incorporated into the design of the Andilaye intervention, these requirements did not go above and beyond what is expected of the HEP [65]. These delivery challenges are consistent with those associated with CLTSH programming and HEP more broadly [38,66]. Additionally, a majority of intervention kebeles had non-active WDALs at the start of implementation, as identified by our Ethiopian-based study team during initial recruitment of activity facilitators. A cross-sectional study in four regions of Ethiopia found similar trends in varying levels of WDAL strategy implementation strength among 423 kebeles [67]. Importantly, findings from Damtew et al. suggest HEP outreach activities were higher in kebeles where active WDAL density was higher (i.e., fewer households per active WDAL).

Although HEWs were paid health workers, WDALs were not. This has brought questions of ethics and sustainability as WDALs are increasingly asked to provide more and more services. Recent qualitative and quantitative studies suggest that unpaid WDALs are actually worse off than their peers and makes women, especially unmarried women, vulnerable to negative gossip and psychological distress [68,69]. Although this point goes much deeper into the political economy, it is an important gap to bring up in the context of empowering women volunteers to enact positive change in their communities [70].

Further, when community health workers are paid to deliver the intervention, there is evidence of
successful delivery [71]. Together, these findings raise questions about the possibility of bringing new programs and approaches to the HEP without adequate support.

Evidence suggests that it is important to move away from information-based interventions to address the array of behavioral factors and determinants that operate at various levels of influence [72–77]. Few sanitation and hygiene interventions employ behavioral theory to locally adapt messages [37,78]. Exceptions include the SuperAmmma intervention, which was developed and implemented in India, and found substantial gains in handwashing with soap [79] and studies that examined the effectiveness of the risk, attitudes, norms, abilities and self-regulation (RANAS) behavioral model to intervention design and showed positive impacts on a variety of WASH behaviors including safe water consumption, solar water disinfection, handwashing, and cleaning of shared sanitation [80–86]. Our intervention aimed to focus on a variety of contextually appropriate behavioral factors rather than knowledge alone. Given the low fidelity of the intervention delivery, further capacity building of federal, regional, and local-level government officials as well as community-level change agents may be necessary for the successful implementation of approaches that move beyond dissemination of information and messages [87–89].

**Study strengths**

Our intervention was theory-informed and included an extensive intervention design process during which we emphasized the solicitation and incorporation of feedback from key stakeholders at regional, zonal, woreda, and community-levels. It was designed to be delivered at scale within the Ethiopia HEP. We utilized a randomized study design, in which intervention and counterfactual communities were allocated to treatment arms randomly. While cluster randomized trials tend to emphasize internal validity, we made considerable effort to enhance external validity. Our study was spread over three woredas in two zones, yielding a heterogeneous mix of contexts. To improve interval validity, we used a
‘fried egg’ [90] approach to select kebeles and minimize spillover. We targeted both rural and peri-urban communities and collected behavioral outcome data on a variety of household members (e.g., primary female caregiver of index child, head of household, all children aged 0-17 years).

Study limitations

The study faced significant delays in gaining local ethical approval to start the project which led to truncated implementation and follow-up periods. Key government actors were less involved than planned, which may have led to sub-optimal fidelity. The integration of Andilaye intervention activities into non-Ethiopian HEP delivery structures (e.g., hired independent community implementers) may yield further investigations into the effectiveness of the intervention on sustained behavior change and mental well-being. Several of our behavioral outcomes were reported (vs. observed), and these types of outcomes may be prone to reporting biases, indicated by differences in our reported and observed measures.

CONCLUSIONS:

We did not find that the Andilaye intervention yielded changes in behaviors and conditions related to sanitation, personal hygiene, or household environmental sanitation; nor did it impact mental health outcomes. Limited integration of Andilaye activities into the HEP likely explains the minimal impact observed and points to considerable challenges related to implementing demand-side interventions at scale in Ethiopia. Evidence from this trial may help address knowledge gaps related to scalable alternatives to CLTSH and inform sanitation and hygiene programming and policy in Ethiopia and beyond. A greater emphasis on implementation research in WASH delivery would support tools and approaches for developing, testing, and adapting scalable best-practice interventions [91].
Acknowledgements

The authors would like to thank our Andilaye study participants, who generously gave their time to participate in our formative work, behavioral trials, and summative assessment surveys. We are grateful for the support we receive from numerous partners at the Democratic Republic of Ethiopia’s Federal Ministry of Health; the Amhara Regional Health Bureau; South Gondar and West Gojam Zonal Health Departments; and the Bahir Dar Zuria, Farta, and Fogera Woreda Health Offices. We would also like to acknowledge support provided by the Health Extension Workers, Women’s Development Army Leaders, and Health Development Army members from our study kebeles. We thank the cadre of field supervisors and enumerators who captured these data (Yeworkwuha Abay, Mantegbosh Abebe, Selamawit Abebe, Tigist Abebe, Rosa Abesha, Mahider Adamu, Balemlaye Addisu, Adanech Admasu, Tirusew Alayu, Adisalem Arega, Yalemwork Asaye, Destaw Asnakew, Yalemwork Ayane, Ayalnesh Belay, Asayech Bimrew, Tigist Bitew, Tiruzer Engidaw, Yeserash Gashaw, Natsenat Gebretsadkan, Woyneshet Genetu, Tewodaj Gizachew, Tibelalech Mihiaret, Yehizbalem Minale, Senait Mulualem, Eleni Nebiyu, Elsabet Seyoum, Mulubirhan Shitu, Sewunet Tadesse, Beza Tesfaye, Rahel Tsegaye, Sintayehu Wasihun, and Maritu Yibrie).


doi:10.1289/EHP7762
Formative research and intervention design

- Formative research, analysis of results, and intervention mapping
- Intervention design feedback from stakeholders and behavioral trials to test intervention approaches
- Finalization of intervention design and materials
- Validation of survey prompts and related answer choices used in the impact evaluation via feedback from formative research households on their meaning, comprehensiveness, and appropriateness
- Intervention allocation at the cluster-level using baseline data to secure balance across key variables

Intervention implementation and process evaluation

Action planning and management activities:
- District-level sensitization and action planning
- Community-level ‘Whole system in the room’
- Ongoing monitoring of implementation
- Adaptive management and cross fertilization visits

Training and capacity building activities:
- Training of trainers and facilitators for counseling visits with caregivers
- Training of community conversations facilitators
- Refresher trainings of trainers and household and group-level intervention activity facilitators
- Ongoing supportive supervision

Key behavior change activities:

- Household-level: counseling visits with caregivers (as dictated by trained facilitators)
- Community-level: mobilization and commitment events
- Group-level: community conversations (as dictated by trained facilitators)

Impact evaluation:
- Enrollment and baseline data collection
- Midline data collection
- Quarterly monitoring visit
- Quarterly monitoring visit
- Endline data collection

Timeline:

<table>
<thead>
<tr>
<th>2016</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sept/Nov</td>
<td>Dec/Feb</td>
<td>Mar/May</td>
<td>Jun/Aug</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>