Evaluation of the Effectiveness of Competency-based Training (CBT) of Women Development Army (WDA) Leaders in Improving the Uptake of Health Care Services in Ethiopia

Evaluation Report

Prepared by:

MERQ Consultancy PLC

Submitted to:

JSI – L10K

Addis Ababa, Ethiopia

Collaborators

IIfPHC, MoH, JSI – L10K, MERQ Consultancy PLC

April 2021

Addis Ababa, Ethiopia

Contributors and Recommended Citation

List of Investigators

Name	Affiliation	Role/Technical Areas
Dr. Frehiwot Nigatu	IIfPHC	Principal Investigator
Yibeltal Kiflie Alemayehu	MERQ	Co-PI, Technical Lead
Dr. Yibeltal Tebekaw	JSI-L10K	Co-PI
Israel Ataro Otoro	MoH	Co-PI
Dr. Mulusew Gerbaba	MERQ	Public Health Expert
Dr. Girmay Medhin	MERQ	Biostatistician
Mekdess Demissie	MERQ	Qualitative Researcher
Bantalem Yeshanew Yihun	JSI-L10K	Community Health
Fasil Walelign Fentaye	MERQ	Program Evaluation
Kassahun Sime	MoH	Community Health

List of Reviewers

Name	Affiliation	Role
Dr. Alula Meressa	MERQ	Reviewer
Challa Tesfaye	MoH	Reviewer
Desalew Emaway	JSI-L10K	Reviewer
Dr. Biruhtesfa Bekele	JSI-L10K	Reviewer
Wuleta Betemariam	JSI-L10K	Reviewer

Recommended Citation

Mulusew Gerbaba, Yibeltal Kiflie Alemayehu, Yibeltal Tebekaw, Israel Ataro Otoro, Girmay Medhin, Bantalem Yeshanew Yihun, Fasil Walelign Fentaye, Kassahun Sime, Alula Meressa, Challa Tesfaye, Desalew Emaway, Biruhtesfa Bekele, Wuleta Betemariam, Frehiwot Nigatu. MERQ Consultancy PLC; JSI-L10K; MoH. 2021.

Acknowledgements

This evaluation was conducted under the auspices of JSI – L10K; we would like to extend our gratitude to JSI – L10K for supporting us both financially and technically. Technical inputs from investigators, co-investigators, and reviewers from the International Institute for Primary Healthcare in Ethiopia, Ministry of Health, and other stakeholders were critical in ensuring the relevance and utility of the evaluation. We would also like to extend our heartfelt thanks to the data collectors, supervisors, and study participants for their participation.

Executive Summary

Background

In 2011, the Ministry of Health (MoH) launched the approach to use Women's Development Army (WDA) structure to promote community engagement and ownership in the health sector. Despite assigning more roles to the WDA, efforts to build their capacity have been limited. Because of this gap and to further expand WDA members' roles, the MoH and its partners started the Competency-Based Training (CBT) program for WDA leaders. The program is a more structured capacity-building initiative from the MoH working with the Federal Technical and Vocational Education Training Institute, regional health bureaus, non-governmental implementing partners, and funding agencies. The program involves training WDA leaders based on Health Extension Program (HEP) packages and communication skills. Those who pass competency assessments receive certification. Evidence on the effectiveness of CBT is not available.

Objective

We evaluated the effectiveness of CBT in improving the knowledge and practices of WDA leaders and their network members. We also identified barriers to and facilitators of implementation.

Methods

We evaluated CBT using a mixed methods design (quantitative with nested qualitative methods) based on Kirkpatrick's model of training evaluation. The model classifies outcomes of training programs into four levels: reaction, learning, behavior, and results. We sampled 561 WDA leaders (of whom 187 received CBT) and 2,376 WDA network members (of whom 911 were from WDA networks with leaders who received CBT) from three regions of Ethiopia, namely Amhara, Oromia, and the Southern Nations, Nationalities, and Peoples Region (SNNPR). In addition, we conducted in-depth interviews with 24 informants, including CBT trainees, trainers, and training coordinators, at woreda and kebele levels to identify barriers to and facilitators of CBT.

Through descriptive analyses, we compared intervention (CBT-trained WDA leaders and their network members) and comparison (nontrained WDA leaders and their network members) groups. We used doubly robust estimation to estimate the average treatment effect on trained WDA leaders and their network members. We used a thematic analysis framework to analyze the qualitative data. We used Stata version 16 for quantitative data analysis and NVivo version 12 for qualitative data analysis.

Findings

Coverage and sociodemographic characteristics of participants: We interviewed 187 trained WDA leaders and 911 network members from the intervention group, and 374 nontrained WDA leaders and 1,825 network members from the comparison group. The response rate ranged from 95.0% to 97.4%. The two groups were largely comparable in their sociodemographic and economic characteristics. However, some background characteristics of WDA leaders differed at a statistically significant level. Compared to nontrained WDA leaders, trained WDA leaders had higher rates of attendance of formal education (65% vs. 45.7%) and membership in community-based health insurance (CBHI; 73.3% vs. 61.8%). Similarly, compared to network members of nontrained WDA leaders, more women among

network members of trained WDA leaders were married (83.5 % vs. 79.5%) and CBHI members (58.1 % vs. 50.6 %), and more had attended school (45.0 % vs. 38.9%).

WDA leaders' satisfaction with CBT: Almost all (98.2%) trained WDA leaders considered CBT useful, and 66.6% were satisfied with the training. Most trained WDA leaders rated training content, mode of delivery, and trainers higher, but felt the timing of the training and the training materials were inadequate. Moreover, respondents rated training duration and the suitability of training days considering the daily routines of WDA leaders unsatisfactory. More importantly, heterogeneity or variations in satisfaction ratings were evident among trained WDA leaders within and between regions.

Knowledge and practices of WDA leaders: Awareness of WDA leaders in relation to water, sanitation, and hygiene (WASH); disease prevention and control; and family health issues was relatively high. However, in-depth knowledge about each topic and practice was generally low and variable. Specific knowledge of toilet use ranged from 48.7% to 88.2% among trained WDA leaders, and 40.9% and 83.7% among nontrained WDA leaders. WASH-related practices were suboptimal despite relatively better practice among trained WDA leaders than nontrained WDA leaders in some of the areas. Open defecation was reported by 16.6% of trained and 23.5% of nontrained WDA leaders.

Access to an improved source of drinking water was high (90.4% among trained and 81.8% among nontrained WDA leaders). Less than half of the WDA leaders in both the trained and nontrained groups had water for handwashing in their compounds, solid waste disposal pits, and liquid waste disposal systems. Similarly, awareness about common communicable and noncommunicable diseases was relatively high in both groups. Nevertheless, some important gaps remained in knowledge required for taking appropriate preventative actions. For instance, 89.9% of trained and 83.1% of nontrained WDA leaders knew that TB can be cured. Knowledge of condom use to reduce the chance of getting HIV was 84.9% among trained and 74.7% among nontrained WDA leaders, yet misconceptions about HIV transmission remained high. Also, knowledge of prevention methods for cervical cancer was suboptimal.

In cases of sickness among household members, 92% of trained and 85.2% of nontrained WDA leaders reported seeking modern healthcare services. Knowledge of family planning (FP) methods was high (94.7% and above) among both trained and nontrained WDA leaders for injectables, implants, and oral contraceptive pills (OCPs). Current use of modern contraceptives was 39.3% among trained and 44.1% among nontrained WDA leaders in the reproductive age group. Coverage of four or more visits for antenatal care (ANC) and health facility delivery was 72% and 55%, respectively, among trained WDA leaders and 58% and 43.3%, respectively, among nontrained WDA leaders. Of trained WDA leaders, 71% reported seeking treatment for a sick child during the last two weeks, while 43% of nontrained WDA leaders reported doing so.

Knowledge and practices of WDA network members: WDA leaders generally reported better knowledge and practices than WDA network members. Knowledge and practices were also mostly better among network members with trained WDA leaders than those with nontrained WDA leaders. Among WDA network members with trained leaders, knowledge of specific benefits of using a toilet ranged from 35.6% to 77.6%, criteria for a healthy home from 18.6% to 65.5%, and methods of preventing insect and rodent infestation from 10.4% to 60.3%. For WDA members with nontrained leaders, knowledge of

specific benefits of using a toilet ranged from 33.4% to 72.7%, criteria for a healthy home from 18.8% to 60.5%, and methods of preventing insect and rodent infestation from 8.5% to 63.9%.

WASH practices were suboptimal among both trained and nontrained WDA leaders. WDA network members with trained WDA leaders reported the following: 20.3% reported open defecation, 27.4% reported the availability of water for handwashing in the compound, and 88.4% reported improved sources of drinking water. WDA network members with nontrained WDA leaders reported the following: 30.5% reported open defecation, 24.7% reported the availability of water for handwashing in the compound, and 82.5% reported improved sources of drinking water.

Awareness about communicable diseases, including TB and HIV, was generally high, and higher among network members with trained leaders than those with nontrained leaders. However, like their leaders, network members' knowledge of detailed issues with implications for prevention was suboptimal. For example, 83.7% of WDA network members with trained and 84.0% with nontrained leaders believed HIV can be transmitted by sharing food with HIV-infected persons. Awareness of the use of ARVs for the prevention of mother-to-child transmission of HIV was 70.2% among WDA network members with trained leaders and 60.5% among members with nontrained leaders. Among network members with trained WDA leaders, 85.1% reported seeking treatment for a sick household member, while 79.6% of network members with nontrained WDA leaders reported doing so.

Knowledge about commonly used FP methods, including injectables, implants, and OCPs, was consistently high among WDA network members in both groups. For the remaining methods, WDA network members with trained leaders were slightly more knowledgeable than those with nontrained leaders. When comparing responses from WDA network members with trained leaders and those with nontrained leaders, 39.7% (vs. 39.4%) reported current use of modern contraceptives, 57.5% (vs. 41.3%) reported at least four ANC visits, and 53.3% (vs. 44.5%) reported health facility delivery. Among households with a sick child in the last two weeks, 49.8% sought treatment for their sick children (54.5% in the intervention group and 47% in the comparison group).

Effect of CBT on HEP-related knowledge and behaviors of WDA leaders and their network members: Doubly robust estimation methods on the effect of CBT on knowledge and behavior of WDA leaders and their network members suggested heterogeneous effects of CBT on different HEP-related knowledge and practices of WDA leaders and their network members. The program improved some of the intended outcomes of CBT for both WDA leaders and network members and some for either the leaders or the members. No evidence suggested that CBT has any effect on WDA leaders or network members for the remaining outcomes.

Among 30 indicators assessed, CBT was positively and significantly associated with 13 (43.3%) for WDA leaders and 19 (63.3%) for WDA members. CBT for WDA leaders were significantly associated with better knowledge about FP, health facility delivery, household treatment of drinking water, covering food and protecting it from flies, improved sources of drinking water, participation in WASH campaigns, knowledge about HIV/AIDS, awareness of NCDs, and awareness of cervical cancer, both among WDA leaders and their network members. On the other hand, for both WDA leaders and their network members, no significant association was revealed between CBT and the use of long-acting reversible

contraception (LARC), postnatal care coverage (PNC), washing hands at critical times, separate kitchen, separate sleeping room, separate room for animals, and improved sanitation facilities.

On average, WDA leaders who received CBT knew 2.4 more FP methods than those who did not receive the training (β = 2.408, 95% CI [1.97, 2.841]). Similarly, the number of FP methods known by WDA members with trained leaders was higher by 1.3 than that among WDA members with nontrained leaders (β = 1.255, 95% CI [1.03, 1.48]). On the other hand, CBT was not significantly associated with the use of modern contraceptives among WDA leaders (β = –0.023, 95% CI [–0.103, 0.056]), while it showed a weak yet statistically significant association with that of WDA members. Network members of trained WDA leaders had 4.8% higher contraceptive prevalence rate (mCPR) than those of nontrained WDA leaders (β = 0.0482, 95% CI [0.012, 0.084]).

Effect of CBT on WDA structure functionality: In the last 12 months, trained WDA leaders conducted a higher number of meetings (M = 7.34) as compared to nontrained WDA leaders (M = 4.92; p = 0.027). The proportion of WDA leaders who conducted at least one meeting in the last three months was higher for the trained WDA leaders (54.1%) than for the nontrained leaders (35.4%; P = 0.01). We also found that the WDA functionality index of WDA networks led by CBT-trained leaders was higher (M = 4.72; SD =2.74) as compared to those led by nontrained ones (mean =1.58; SD = 2.26; P = 0.001). After adjusting for potential confounders, WDA networks led by trained leaders have a 2.555 higher WDA functionality index than those led by nontrained leaders (β =-2.255, 95% CI [1.898, 2.612]).

Facilitators and barriers of CBT implementation: The implementation of CBT was facilitated and hindered by factors operating at individual, community, and health-system levels. Factors such as trainers' motivation, trainees' level of understanding or literacy level, trainer–trainee relationships, community acceptance, distance from the training place, availability of resources such as teaching materials, incentives, and coordination emerged as important determinants of the implementation of CBT.

Conclusions and recommendations

WDA leaders who participated in CBT found the training useful and its processes somewhat satisfactory. CBT was effective in improving the knowledge and practices of WDA leaders and their network members in relation to WASH, disease prevention and control (DPC), and maternal and child health (MCH). The effect of CBT was, however, not uniform across all targeted behaviors. CBT had no or a negligible effect on several health-related behaviors identified as objectives of the training program.

The implementation of CBT was facilitated and hindered by factors operating at individual, community, and health-system levels. Findings of this evaluation have important health policy and programmatic implications in relation to building the capacity of voluntary community health workers in Ethiopia. Therefore, we provide recommendations to address key findings of the evaluation. The observed heterogeneity between and within regions in the implementation of CBT and satisfaction levels of trainees calls for reexamining the WDA selection criteria, teaching methods, training materials, and refreshments and addressing context-specific barriers such venue and timing. The variability in the effectiveness of CBT in achieving its objectives, including the lack of effect on some of the objectives, calls for more stringent monitoring of implementation fidelity and the design of mechanisms to address

additional determinants of health behavior (social determinants). Finally, addressing the multi-level determinants of CBT implementation implies the need to consider the involvement of other stakeholders, leadership commitment, and other viable approaches to strengthen the capacity of community health volunteers in addition to strengthening the current training.

Table of Contents

Contributors and Recommended Citation i
Acknowledgements ii
Executive Summary iii
Lists of Figures xi
List of Tables xii
Acronyms xiii
1 Background 1
2 Description of competency-based Training 2
3 Evaluation Questions and Objectives 3
3.1 Evaluation questions 3
3.2 Objectives of the evaluation 4
4 Evaluation Methods and Procedures 4
4.1 Study setting 4
4.2 Conceptual framework 4
4.3 Evaluation design 5
4.4 Populations and sampling 6
4.4.1 Source population 6
4.4.2 Study population 6
4.4.3 Sample size determination 6
4.4.4 Sampling strategy 7
4.5 Measurements 7
4.6 Data collection tools 10
4.7 Data collection teams and training 10
4.8 Pretesting of data collection tools 10
4.9 Data collection 11
4.10 Data quality assurance procedures 11
4.11 Data management and analysis 11
4.11.1 Quantitative data management and analysis 11
4.11.2 Qualitative data management and analysis 12
4.12 Ethical considerations 13
5 Findings 14
5.1 Coverage and sociodemographic characteristics of participants 14
5.1.1 Study coverage and response rate 14
5.1.2 Background characteristics 14
5.2 Reaction: Satisfaction of WDA leaders with CBT 16
5.2.1 CBT experiences of trained WDA leaders 16
5.2.2 WDA leaders' satisfaction with CBT 18
5.3 Knowledge and practices of WDA leaders 24
5.3.1 Water, sanitation, and hygiene 24
5.3.2 Disease prevention and control 26

5.3.3 Family Health 28
5.4 Knowledge and practices of WDA network members 34
5.4.1 Water, sanitation, and hygiene 34
5.4.2 Disease prevention and control 36
5.4.3 Family health 37
5.5 Effect of CBT on HEP-related knowledge and behaviors among WDA leaders and their network members 43
5.6 The effect of CBT on the functionality of WDA structures 45
5.7 Facilitators of and barriers to CBT 49
5.7.1 Performance of CBT in study woredas 49
5.7.2 Facilitators of and barriers to the performance of CBT 49
5.7.3 Macro-level factors operating at health system policy and practices 52
6 Discussion 57
6.1 Discussion of key findings 57
6.2 Strengths and limitations of the study 60
7 Conclusion and Recommendations 61
7.1 Conclusion 61
7.2 Recommendations 61
References 64
List of Annexes 65
Annex 1: Listing and sampling formats 65
Annex 2: Data collection tools 65
Annex 3: List of study woredas: Initial sample and replaced woredas 65

Lists of Figures

Figure 1. CBT activities and expected outcomes (based on Kirkpatrick's model of training evaluation) 5

Figure 2. Graphical description of evaluation design 5

Figure 3. Qualitative data management and analysis processes 13

Figure 4. WDA leaders' reactions regarding the usefulness of CBT, February 2021 18

Figure 5. Satisfaction of trained WDA leaders with CBT program, February 2021 19

Figure 6. WDA leaders' satisfaction with CBT program by age group, February 2021 19

Figure 7. WDA leaders' reactions to CBT processes, February 2021 20

Figure 8. WDA leaders' reactions to CBT module contents, February 2021 22

Figure 9. Knowledge of FP methods in trained and nontrained WDA leaders, February 2021 29

Figure 10. Contraceptive method mix among current users of FP among trained and nontrained WDA leaders 29

Figure 11. Awareness about danger signs during pregnancy among trained and nontrained WDA leaders 31

Figure 12. Awareness about danger signs during pregnancy among trained and nontrained WDA leaders by CBT status 31

Figure 13. WDA leaders' knowledge about timing of first PNC visit, February 2021 32

Figure 14. Vaccination status among children of trained and nontrained WDA leaders (12–23 months age), February 2021 33

Figure 15. Knowledge of FP methods among WDA network members by group 38

Figure 16. Contraceptive method mix among current FP user WDA network members with trained and nontrained leaders 40

Figure 17. Awareness about danger signs of pregnancy among WDA network members with trained and nontrained leaders, February 2021 40

Figure 18. Vaccination coverage among children 12–23 months of WDA members with trained and nontrained leaders, February 2021 42

Figure 19. Childhood illnesses and treatment-seeking behaviors among WDA network members with trained and nontrained leaders, February 2021 42

Figure 20. Indicators of aspects of CBT that affected WDA leaders and network members 43

Figure 21. WDA implementation functionality score by region, February 2021 47

Figure 22. Implementation status of CBT activities among sample kebeles by region 49

Figure 23. Categories of barriers to and facilitators of CBT implementation, February 2021 50

List of Tables

Table 1. Competency-based Training modules and descriptions of topics 2
Table 2. Descriptions of outcome levels (Kirkpatrick's model of training evaluation) 4
Table 3. Background characteristics of WDA leaders stratified by CBT participation, February 2021 15
Table 4. Duration of training, assessment, and certification status of CBT-trained WDA leaders, February 2021 16
Table 5. Summary statistics of CBT-trained WDA leaders' reactions to CBT, February 2021 21
Table 6. WASH facilities and knowledge status among CBT-trained and nontrained WDA leaders, February 2021 24
Table 7. Practices related to WASH among CBT-trained and nontrained WDA leaders, February 2021 25
Table 8. Knowledge about TB and HIV/AIDS among CBT-trained and nontrained WDA leaders, February 2021 27
Table 9. Knowledge about NCDs, cervical cancer, adulthood illness, and treatment-seeking behaviors among CBT-trained and nontrained WDA leaders, February 2021 28
Table 10. FP practices among WDA leaders by training status, February 2021 30
Table 11. Knowledge and practices related to maternal and child healthcare services among WDA leaders, February 2021 32
Table 12. WASH-related knowledge among WDA network members with trained and nontrained leaders, February 2021 34
Table 13. WASH-related practices of WDA network members, February 2021 35
Table 14. Awareness of, knowledge of, and practices related to TB and HIV/AIDS among WDA network members with trained and nontrained leaders, February 2021 36
Table 15. Awareness of, knowledge of, and practices related to NCDs, adulthood illnesses, and treatment-seeking behaviors among WDA network members with trained and nontrained WDA leaders, February 2021 37
Table 16. Reproductive history and use of FP methods among WDA network members with trained and nontrained leaders, February 2021 39
Table 17. Use of maternal healthcare services among WDA network members with trained and nontrained leaders, February 2021 41
Table 18. Effect of CBT on knowledge and practices of WDA leaders and network members, February 2021. 44
Table 19. Association of training hours, competency assessment, and certification with health outcomes 45
Table 20. Functionality of WDA structures led by trained and nontrained WDA leaders, February 2021 46
Table 21. Summary of most frequently reported barriers to and facilitators of CBT, by level 57

Acronyms

Acronyms	Description
ANC	Antenatal care
ATE	Average time effect
CBT	Competency-based training
CoC	Certificate of competency
DC	Delivery care
DPC	Disease prevention and control
EPHA	Ethiopian Public Health Association
FP	Family planning
HIV/AIDS	Human immune virus/acquired immunodeficiency syndrome
HEP	Health extension program
HEW	Health Extension workers
IRB	Institutional review board
IUD	Intrauterine device
LMICS	Low- and middle-income countries
MoH	Ministry of Health
MoE	Ministry of Eduction
NCD	Noncommunicable diseases
OCP	Oral contraceptive pill
PNC	Postnatal care
SNNPR	Southern Nations, Nationalities, and Peoples Region
TB	Tuberculosis
WASH	Water, sanitation, and hygiene
WoHO	Woreda Health Office
WDA	Women's Development Army

1 Background

Ensuring universal access to primary healthcare services has been a major challenge for low- and middle-income countries (LMICs), and Ethiopia is no exception (Bitton et al., 2019). In 2003, the Government of Ethiopia launched the Health Extension Program (HEP) to extend essential healthcare services to the community and household levels (Bilal et al., 2011). The HEP involves the delivery of packages of health promotion, disease prevention, and limited basic curative services to individuals and families. Primarily health extension workers (HEWs) are responsible for the implementation of HEP. They have been providing healthcare services at home, at health post, and in community settings, and they also visit and provide a bridging service for communities with health centers (FMoH, 2008).

Community engagement and ownership is a core principle of HEP and an important strategic objective of the Ethiopian health sector (Bilal et al., 2011; Bilal, 2012). The role of communities in HEP extends to promoting service delivery, as they are involved in demand-creation activities through dissemination of health information. For instance, the Women's Development Army (WDA) approach was used in the health sector and linked with HEP to promote community engagement and ownership (Banteyerga, 2011). WDA networks are networks of six household members connected structurally with 1 to 5 networks in the same neighborhood. A group of approximately six 1-to-5 leaders (hereafter, 1-5 leaders) is in turn led by a "1-to-30 leader" (hereafter 1-30 leader). The 1-to-30 leaders serve under the direct supervision of a HEW and serve as volunteer community health workers. They are supposed to help to educate the 1-5 leaders and members on topics such as seeking antenatal and birth care, hygiene, and sanitation-related behaviors and to organize activities such as immunization campaigns (Assefa et al., 2019). WDA leaders are expected to have greater knowledge of HEP topics and to perform well in implementing HEP, have good communication skills, be active in community development, and be ready to teach others and influence the behavior of their respective network members (Maes et al., 2018; Yitbarek et al., 2019). WDA leaders demonstrate these qualities through the dissemination of health information, demonstration of model health behaviors, and mobilization of community members in the planning, implementation, monitoring, and evaluation of healthcare services at the community level (Yitbarek et al., 2019; Damtew et al., 2018).

Previous studies revealed that the WDA has improved the outcomes of maternal and child healthcare services such as skilled delivery and antenatal care (ANC), maternal death, and child immunization (Yitbarek et al., 2019). Yet despite their increasing roles and responsibilities in the provision of community-based healthcare services, the 2019 national assessment of the HEP indicated that WDA leaders do not practice the desired health behaviors, are not widely accepted by community members, and have low levels of motivation to serve as voluntary community health workers (Teklu et al., 2020; MERQ Consultancy PLC, 2019). Moreover, efforts to build their capacities have been limited (Damtew et al., 2018). These efforts have included Competency-based Training (CBT), a structured capacity-building initiative targeting WDA leaders.

CBT was implemented in 2018 through a collaboration between the Ministry of Health (MoH), the Federal Technical and Vocational Education and Training Institute (TVET), the Ministry of Education

(MoE), regional health bureaus (RHBs), and non-governmental implementing partners and funding agencies. So far, close to half a million WDA leaders have been trained through this program. Understanding the extent to which CBT improves the capacity of WDA leaders to change their own and their network members' knowledge, attitude, skills, and practices is critical before investing more resources into further scaleup of the training. This study, therefore, intended to determine the effectiveness of CBT in improving the knowledge, skills, and practices of WDA leaders and their network members. The findings will inform the design and implementation of future capacity-building efforts targeting voluntary community health workers.

2 Description of competency-based Training

CBT involves training WDA leaders on health-related topics, based on HEP packages, and communication skills. Those who pass competency assessment receive certification. The training uses a curriculum that intends to improve the knowledge, skills, and attitudes of WDA leaders and reach the general population through cascading training and the diffusion of exemplary behaviors. Providing WDA leaders with certificates of competency (CoCs) is also expected to improve self-esteem and thus their willingness to serve their communities.

The training manual was developed in a series of steps that engaged relevant stakeholders. First, the curriculum and content of the training materials were developed by the MoH, MoE, and TVET working team based on the available literature and experience from relevant projects. The training materials were translated into local languages (Amharic, Afan Oromo, and Tigrigna) before a team of local experts reviewed them to ensure they were culturally adapted to the local context. The training materials were organized into seven modules: (1) improving teamwork and development; (2) basic personal and environmental hygiene; (3) preventing major communicable, noncommunicable, and neglected tropical diseases; (4) family planning (FP); (5) ANC, (6) postnatal care (PNC) and child growth; and (7) life-saving services (see Table 1).

Table 1. Competency-based Training modules and descriptions of topics

Session	Title	Description (topics and module length)
Module 1	Improving teamwork and development	Team formation, steps for conducting a meeting, household/community mapping and planning, and use of reporting forms; 8 hours
Module 2	Use of basic personal and environmental hygiene	Personal hygiene, types of improved toilets, handwashing, water treatment, and vector control; 10 hours
Module 3	Prevention of major communicable, noncommunicable, and neglected tropical diseases	Communicable and noncommunicable diseases; 6 hours Communicable diseases: signs and symptoms of malaria, TB, and HIV/AIDS, and prevention and control mechanisms NCDs: hypertension, diabetes mellitus, mental illness, cardiovascular diseases, and asthma Tropical diseases: trachoma, onchocerciasis, guinea worms, leishmaniasis, podoconiosis, and scabies

Session	Title	Description (topics and module length)
Module 4	Family planning	FP or contraceptive methods, use, and adolescent health problems; 4 hours
Module 5	Antenatal care	Danger signs and complications of pregnancy and delivery, use of ANC and delivery care (DC) at healthcare facilities, harmful traditions, and nutrition during pregnancy; 6 hours
Module 6	Postnatal care	PNC services and timing, danger signs, benefits of breastfeeding, and referral system; 4 hours
Module 7	Life-saving strategies: child growth, development, and survival	Neonatal health, child growth and development, breastfeeding, complementary feeding preparation, causes of infant death, vaccinations, and infant feeding; 14 hours

These modules are completed in 52 hours, with each module having specified hours ranging from 4 to 14 hours. After the trainees complete each module, they sit for a competency assessment organized at the kebele level. They should pass the competency assessment to become certified. If a trainee fails the competency assessment, she must undergo additional training and reassessment. As part of the continuous capacity-building for WDA leaders, annual refresher training is organized with a focus on packages that need more engagement.

Since its introduction in 2018, 16% (half a million) of the 3 million WDA leaders the MoH plans to train have been trained on level I. Of those who completed the training, 108,858 (36.3%) completed competency tests administered under the oversight of competency testing agencies, and 99,702 (91.6%) received certificates of competency (CoCs) after passing.

3 Evaluation Questions and Objectives

CBT was designed to increase the capacity of WDA leaders in terms of knowledge, health behavior, and the skills needed to communicate with and influence their network members. The MoH sought to evaluate the effectiveness of the training to inform decisions related to scaling up the program and to extract lessons on capacity-building for community volunteers in general.

3.1 Evaluation questions

The assessment was intended to answer the following questions:

- How did WDA leaders react to, or how satisfied were they with, CBT?
- Did CBT improve the knowledge of the primary targets, WDA leaders?
- Did CBT improve adoption of HEP-related behaviors among trained WDA leaders?
- Did the participation of WDA leaders in CBT lead to better health knowledge and behaviors among the network members of trained WDA leaders?

3.2 Objectives of the evaluation

To answer the above questions, the evaluation addressed the following objectives:

- Assess WDA leaders' reactions to CBT
- Compare the knowledge and practices (implementation of HEP packages) of trained and nontrained WDA leaders and their network members
- Estimate the effects of CBT on the knowledge and practices of trained and nontrained WDA leaders and their network members on selected health outcomes.
- Identify barriers to and facilitators of capacity-building for WDA leaders through CBT

4 Evaluation Methods and Procedures

4.1 Study setting

The evaluation was conducted in three regions of Ethiopia—Amhara, Oromia, and the Southern Nations, Nationalities, and Peoples Region (SNNPR). CBT was provided to WDA leaders by HEWs working in rural kebeles. The study was conducted in rural kebeles where WDA leaders completed CBT 12–18 months before. The study period was three months (January to March 2021).

4.2 Conceptual framework

We evaluated the effectiveness of CBT based on Kirkpatrick's model for the evaluation of training programs. Kirkpatrick's model has been widely and flexibly used to assess the effectiveness of training programs (La Duke, 2017). The model identifies training program outcomes that range from the direct reaction of training participants to technical and nontechnical aspects of training to the achievement of the results the program intends to achieve. The model classifies these results into four levels—reaction, learning, behavior, and results (Bates and Planning, 2004; La Duke, 2017). Descriptions of the four levels of training outcomes, according to Kirkpatrick's model of training evaluation, are presented in Table 2.

Table 2. Descriptions of outcome levels (Kirkpatrick's model of training evaluation)

Level	Description
Reaction	First level used to understand participants' satisfaction with the training process. Responses of trainees to their experience during training and level to which training participants are satisfied with various aspects of the training.
Learning	Level of outcome that focuses on changes in knowledge among training participants that can be attributed to the training.
Behaviors	Measure helps to clarify whether training participants have been applying what they learned in their daily practice what they learned. Measures trainees' change in behavior (whether learning is translated into practice).
Results	The highest level of outcomes in the model focusing on the achievement of higher-level changes among targets of trainees, including organizations in which the trainees work, communities where trainees live, and any other setting the training aims to influence through trainees.

For the CBT program, reaction, learning, and behavior are outcomes observed among WDA leaders who receive the training. We expected to observe results in the form of the functionality of WDA structures and in terms of improved knowledge and health behaviors among network members (Figure 1)

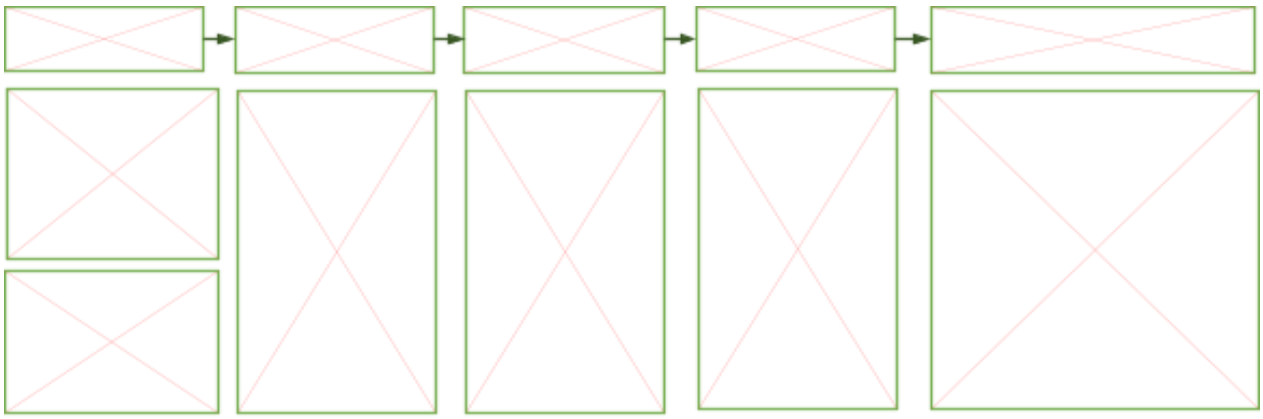


*Figure SEQ Figure * ARABIC 1. CBT activities and expected outcomes (based on Kirkpatrick's model of training evaluation)*

4.3 Evaluation design

We employed a mixed methods evaluation involving quantitative and qualitative methods of data collection. We generated quantitative evidence using posttest-only comparison group design (Krishnan, 2019) with the intention to examine if and to what extent participation of WDA leaders in CBT positively influences relevant outcomes (HEP-related knowledge and practices among WDA leaders and their network members; see Figure 2).

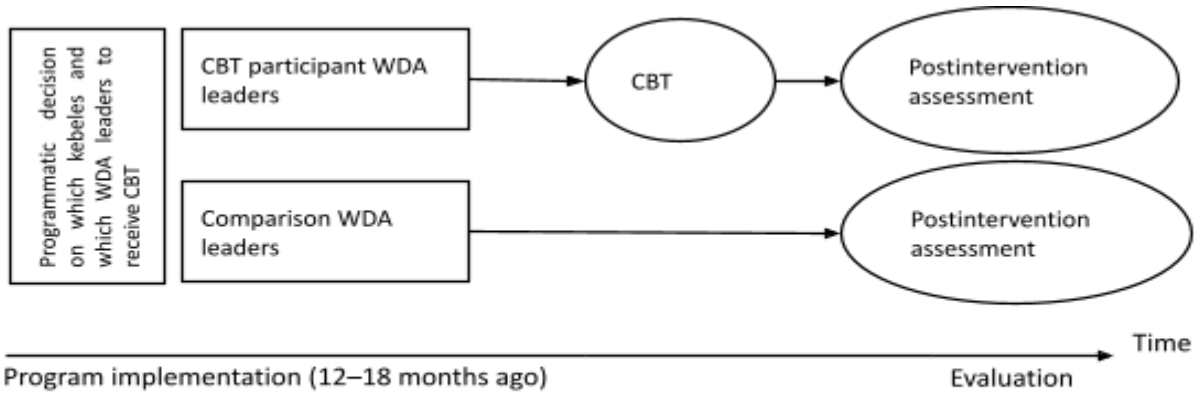


*Figure SEQ Figure * ARABIC 2. Graphical description of evaluation design*

WDA leaders who participated in the CBT program 12–18 months ago and their network members were considered as the intervention group. WDA leaders who did not participate in CBT and their network members were used as the comparison group. Both groups were included in the postintervention assessment to investigate the relationships between participation in CBT and outcomes of interest. The aim was to use the mean scores on the posttest to examine if the CBT program affected outcomes of

interest for those who attended the program. In addition, qualitative data were collected from WDA leaders, trainers, and training coordinators to complement quantitative findings.

4.4 Populations and sampling

4.4.1 Source population

We used two source populations to recruit study participants: the WDA leaders and their network members. To obtain both intervention and comparison groups from similar woredas, we included woredas with partial implementation of CBT (woredas where CBT was implemented in 25% to 75% of their kebeles in the last 12–18 months).

4.4.2 Study population

We considered WDA leaders randomly selected from a sampling frame obtained from randomly selected kebeles as study participants. For each WDA leader selected, we included her network members as study participants. We selected trained WDA leaders from kebeles where CBT was implemented; we selected comparison WDA leaders from kebeles where CBT had not been implemented. We did this to avoid contamination. We chose CBT-implementing woredas using the criteria that the woreda should have (a) a 25% to 75% CBT implementation rate of total health posts in the last 12–18 months and (b) at least five kebeles eligible for the intervention group (have implemented CBT) and at least five kebeles eligible for the comparison group (have not implemented CBT). We used these criteria only to identify woredas from which we could select both intervention and comparison groups.

For WDA leaders and their network members to qualify for inclusion in the intervention group, WDA leaders should have completed CBT (as reported by HEWs) 12–18 months before the study and should be working as a WDA leader. In the comparison kebeles, all WDA leaders reported by HEWs were eligible for the study. In both groups, if we selected a WDA leader for the study, her network members were automatically included.

4.4.3 Sample size determination

The sample size was estimated using a sample size calculation formula for two population proportions based on the following assumptions:

- Given the time between the provision of CBT and the evaluation (12–18 months), we assumed that outcomes related to hygiene and environmental sanitation improved because of the program. Therefore, we calculated sample size based on the proportion of WDA leaders with improved sanitation, assumed to be 20%, and hand-washing facilities available, assumed to be 6.7%, with these assumptions based on the 2019 national HEP assessment survey. We took the maximum of the sample sizes calculated based on these proportions.
- CBT of WDA leaders will increase the desired behavior by 17%.
- Design effect of 1.5 because of multistage sampling (the three regions as strata, woreda as first-stage sampling unit, kebele as second-stage sampling unit, and WDA leader as third-stage sampling unit) and 90% power to detect the difference between CBT recipients and nonrecipients.
- Nonresponse rate expected not to exceed 6%.

Accordingly, the final sample size was estimated to be 192 WDA leaders along with approximately 960 network members in the intervention group, and 384 WDA leaders and 1,920 network members in the comparison group.

4.4.4 Sampling strategy

We used a three-stage sampling technique to recruit study participants from the three regions, namely Oromia, Amhara, and SNNPR. In the first stage, we sampled woredas within each region, resulting in eight woredas sampled. We obtained a sampling frame of woredas, along with their status in implementing CBT, from RHBs. Subsequently, we created lists of woredas that had delivered CBT in 25–75% of their kebeles, and then we randomly selected three woredas from Oromia, three woredas from Amhara, and two woredas from SNNPR. We replaced woredas we found to be ineligible upon verification from Woreda Health Offices (WoHOs), because of timing of training or coverage of training, with eligible woredas from a similar zone. The selected woredas are listed in Annex 3.

In the second stage, we sampled kebeles from the sampled woredas. Kebeles in each selected woreda were stratified into intervention and comparison kebeles. Within each woreda, we randomly selected four kebeles among those that provided and implemented CBT for more than 12 months (intervention group) and four kebeles among those that had not provided CBT (comparison group). In the third stage, we selected WDA leaders from the selected kebeles. We obtained lists of WDA leaders (separate lists for 1-5 - and 1-30 WDA leaders) from HEWs and used the lists as a kebele-level sampling frame. Using the Excel RAND command, we randomly selected the intervention group, which included six WDA leaders (five 1-5 WDA leaders and one 1-30 WDA leader), and the comparison group, which included 12 WDA leaders (ten 1-5 and two 1-30 WDA leaders). For each WDA leader selected, we approached her five WDA network members for consent and then interviewed them. If the network group comprised more than five members, we randomly selected five network members. The detailed sampling procedure is set out in Annex 1.

We selected key informants for qualitative data collection based on roles they had in the CBT program. From each region, we purposively selected trained WDA leaders, HEWs who participated as trainers, and HEP coordinators or other experts from WoHOs or health centers who participated as local coordinators of the CBT program.

4.5 Measurements

The measurements were developed based on Kirkpatrick's model, which identifies four levels of outcomes for training programs: reaction, learning, behavior, and results.

Reaction: We assessed the reactions of WDA leaders to CBT using 25 items, each recorded on a 5-point Likert scale (0: strongly disagree, 1: disagree, 2: neutral, 3: agree, 4: strongly agree). These items address different domains: relevance of topics covered, appropriateness of training methods, clarity of content, convenience of setup (time, place, language), and overall levels of satisfaction. We then checked the distribution of each item using both exploratory factor analysis and principal component analysis. We removed items with a factor loading less than 0.4 and then computed both mean score (ranging from 0 to 4) and sum of score of items (ranging from 0 to 100). We also computed the predicted values of the

satisfaction score and checked them by examining histograms and box plots. We also descriptively summarized the regional differences in satisfaction score values.

Learning: We assessed the effectiveness of CBT in terms of learning by estimating differences in the knowledge levels of trained and nontrained WDA leaders. In this assessment, we used items that measured participants' knowledge of WASH, ANC, delivery care, PNC, child immunization, and disease prevention and control. For each package (*i.e.*, CBT module), we developed specific items related to teamwork, family health (including FP, ANC, PNC), disease prevention and control, hygiene and environmental sanitation, and life-saving strategies.

Behavior: We assessed trained and nontrained WDA leaders' levels of adoption of HEP-related behaviors and communications with network members on different HEP-related topics. We measured use of healthcare services, including family healthcare (FP, ANC, DC, PNC) practices, disease prevention and control, hygiene and environmental sanitation, and treatment-seeking behaviors to determine the effectiveness of CBT in terms of achieving behavior change among primary targets of the training.

Results: We measured results in terms of the possible influence of WDA leaders on their network members. Therefore, we assessed the same measures of knowledge and behavior applied to WDA leaders to measure learning and behavior among network members to evaluate results.

Exposure to CBT: We considered a WDA leader exposed to CBT if a HEW reported that she had attended CBT and the WDA leader verified this herself. We considered all network members exposed to CBT if their WDA leader was trained. If the WDA leader did not participate in CBT at all, we considered her and her network members unexposed.

Specific measures of outcomes

We considered the following outcomes to estimate the effect of CBT on WASH, maternal and child healthcare services and disease prevention and control (DPC), and WDA functionality.

Family planning use: We asked respondents about the use of any FP (modern or traditional) methods. We treated their responses as dichotomous variables coded 1 if women of reproductive age (15–49 years) or their partners were using any contraceptive method (*e.g.*, OCP, injectable, condom, natural method, implant, IUD, and female or male sterilization) and 0 otherwise.

Knowledge of contraceptive methods: We measured this as the percentage of respondents, of currently married women, and of women who know any contraceptive method, by specific method. We then aggregated this variable in terms of number of methods known by each respondent by counting the number of methods a woman knows.

ANC: We measured ANC using a binary response item that asks respondents whether they received ANC services from a health facility or not during the last 12 months prior to the survey.

Delivery services: At the time of the survey women were asked whether they delivered their most recent child, born in the last 12 months, at a health facility (hospital, health center, health posts). We created a binary response variable and coded institutional delivery as 1 and other as 0.

Postnatal care: We asked all women who gave birth during the 12 months preceding the survey, "Did you have a PNC visit for your youngest child, which you delivered in the last 12 months?" We coded their responses 1 if they received PNC and 0 otherwise.

All basic vaccinations: We measured child vaccination coverage among children aged between 12 and 23 months. We collected information on immunization from the children's immunization cards and interviews with their mothers. We considered children who had already taken a dose of bacille Calmette-Guerin (BCG), three doses of pentavalent (DPT-HepB-Hib), three doses of the polio vaccine, and a dose of the measles vaccine fully vaccinated and coded 1. We considered the rest not fully vaccinated and coded 0.

WDA structure functionality index: We used nine item instruments to assess WDA functionality. Questions in the instrument ranged from asking whether the WDA leaders conducted meetings, assisted HEWs, planned their activities, implemented their plans (assisted, conducted home visits, disseminated health messages), and reported their performance at kebele levels. Each of these items is binary and categorical. We created WDA implementation intensity scores with values ranging from 0 to 9. This score was generated by adding responses to the nine items intended to measure the proposed activities of the WDA.

Knowledge of sanitation and hygiene: We measured knowledge related to sanitation and hygiene using sets of items that address a healthy environment, the prevention of human contact with waste, safe waste practices, and liquid disposal. We created a sum-of-knowledge score from these items.

Handwashing facilities: We measured the presence of handwashing facilities by direct observation and by asking respondents whether they practiced a set of hygiene and sanitation practices for the preservation of health and healthy living. If the observer did not see a specific location for handwashing in the house, yard, or plot during data collection, we considered households to have no handwashing facilities.

Handwashing at critical times: We asked respondents whether they washed their hands before preparing food, before eating, after eating, after cleaning babies, and after visiting the toilet. We coded "yes" responses 1 and "no" responses 0.

Aware of tuberculosis, noncommunicable diseases, and cancer: We also asked respondents whether they had heard of TB, NCDs, and cancer. We coded "yes" responses 1 and "no" responses 0.

Other covariates

We identified potential covariates based on a priori. We retrieved covariates associated with WASH and maternal healthcare services from literature and documented them for use in multivariable analysis to adjust for potential confounding effects. We identified three categories of variables: individual-, household-, and community-level variables. Individual- and household-specific characteristics include the age of women, education, and the household wealth index. The household wealth index is a composite variable based on households' ownership of 21 items, with data collected in the HEP surveys. This index categorizes women into five categories (poorest, poorer, middle, richer, and richest).

4.6 Data collection tools

The study team drafted data collection tools for the household survey and key informant interviews. The team used previously tested standard items for household surveys, mainly questions from Demographic and Health Surveys, to develop the quantitative data collection tool. We presented the draft tools to stakeholders for review and amended the tools based on recommendations from the review.

Each of the data collection tools was translated to local languages (Afan, Oromo, and Amharic). The household survey questionnaire was programed for electronic data collection using a CSPro data collection template. We prepared qualitative data collection guides with open-ended questions and probing items with a focus on the four levels of outcomes, and barriers to and facilitators of CBT.

4.7 Data collection teams and training

We recruited 48 data collectors and eight supervisors based on their experience with similar community-based studies, gender, and ability to speak local languages in the study areas. We then provided a five-day training course, including pretesting of data collection tools and procedures. The training content briefly covered the following:

- Background information about CBT evaluation
- Data collection instruments—item-by-item discussion
- Field data collection procedures, including listing of sampling units at various levels, sampling of kebeles within study woredas, sampling WDA leaders within sampled kebeles, and identification of WDA network members for selected WDA leaders
- Procedures for interviews, including ethics and the consent process, and tablet management for field data collection and data submission

The training involved lectures, group discussions, and role play. We used daily training evaluations (written and oral feedback) to monitor and improve the training quality. Training facilitators reviewed the daily reflections and presented summaries of the reflections and proposed actions during the first session of the next day. At the end, we grouped data collectors and supervisors in eight teams composed of one supervisor and six data collectors for each woreda.

4.8 Pretesting of data collection tools

We chose Ada'a woreda, which has 24 kebeles, for the pretest because the woreda has implemented CBT and is located at a manageable distance from Addis Ababa. The pretest was done on the fourth day of the data collection teams' training. We organized the data collection team into two pretest kebeles. The team interviewed ten WDA leaders and ten network members. Both teams collected all comments raised about the tools, translation, and CSPro programming. We discussed these comments on the fifth day of the data collectors' training and incorporated them in the tools.

4.9 Data collection

We collected quantitative data through face-to-face interviews with WDA leaders and their network members using a pretested structured questionnaire. We collected qualitative data through key informant interviews with CBT trainees, trainers, and training coordinators. Data collection teams started their field work activities after receiving letters written to the woredas in their respective regions. The supervisors and the research team supervised the data collection process and provided daily feedback on the completeness and consistency of data.

The researchers followed up on challenging issues related to tools, data collection procedures, identification of replacement woredas, and the sampling of WDA leaders and their network members. The supervisors were responsible for the key informant interviews, and they recorded detailed notes of the narratives. Data were transferred instantly from the field to the central data management unit at MERQ, and subsequently extracted and submitted to the survey team using the appropriate CSV format.

4.10 Data quality assurance procedures

We employed several measures at each stage of the survey. We used a questionnaire with items adapted from previous studies in Ethiopia. We had all tools translated into local languages and checked for consistency. We recruited experienced data collectors from MERQ's database of data collectors and supervisors, stored according to experience and performance in data collection activities. We provided data collectors with interactive training for five days. We also used CSPro with validated skip patterns, whereby wrong or out-of-range answers are rejected. The research team followed the progress of field data collectors daily using phone calls.

In addition to telephone-based follow-up, a team from MoH supervised two woredas on site and provided feedback on the data collection process. We finally cleaned the data in a separate file using STATA to prepare it for the final analysis. We checked for outliers and missing values. Thus, we followed proper data management procedures. We also examined the distribution of variables before choosing identification or estimation strategies.

4.11 Data management and analysis

4.11.1 Quantitative data management and analysis

We used descriptive analysis such as mean difference and employed multivariable analyses to examine the effect of CBT on WDA leaders' and their network members' knowledge of and practices related to HEP services. First, we estimated ordinary least-squares regressions, adjusting for potential confounding covariates. These covariates included age, education status, partner's education status, wealth index of households, agricultural land and livestock ownership, marital status, and reproductive history.

No causal interpretation can be associated with such analysis due to unobserved factors that may result in systematic differences between those who received CBT and those who did not. Therefore, we tried to generate unbiased estimates by employing a doubly robust estimation technique that combines inverse

probability weighting and regression adjustment or ordinary least squares estimation (Wooldridge, 2010) to estimate the causal effect of CBT on outcomes of interest.

The doubly robust method correctly specifies two regression models for the outcome and the treatment as a function of the covariates. The propensity score-based weighting model tested the probability that a WDA leader received the treatment (*i.e.*, CBT) given a function of their covariates or baseline characteristics. The ordinary least-squares model predicted the outcome as a function of the covariates among the exposed (*i.e.*, received CBT) and unexposed groups (*i.e.*, WDA leaders who did not receive CBT).

In this study, we aimed to estimate the effects of CBT on the outcomes of interest based on counterfactual outcomes; that is, what would have happened to study participants if they had not participated in CBT interventions. Hence, in all models, we reported the effects of CBT among those who received the training; that is, the average treatment effect among the treated (Funk et al., 2011). We used STATA version 16 for all statistical analyses. We considered a p-value of < 0.05 with a 95% confidence interval (CI) as statistically significant.

4.11.2 Qualitative data management and analysis

Audio records from the 24 informants were transcribed and translated into English. We then analyzed the data using a thematic analysis framework assisted by NVivo version 12 software. First, the qualitative research team familiarized themselves with the transcripts. We then developed a draft-coding framework based on the study objectives, content of data collection guides, and content of transcripts. Details of the qualitative data management and analysis process are described in Figure 3.

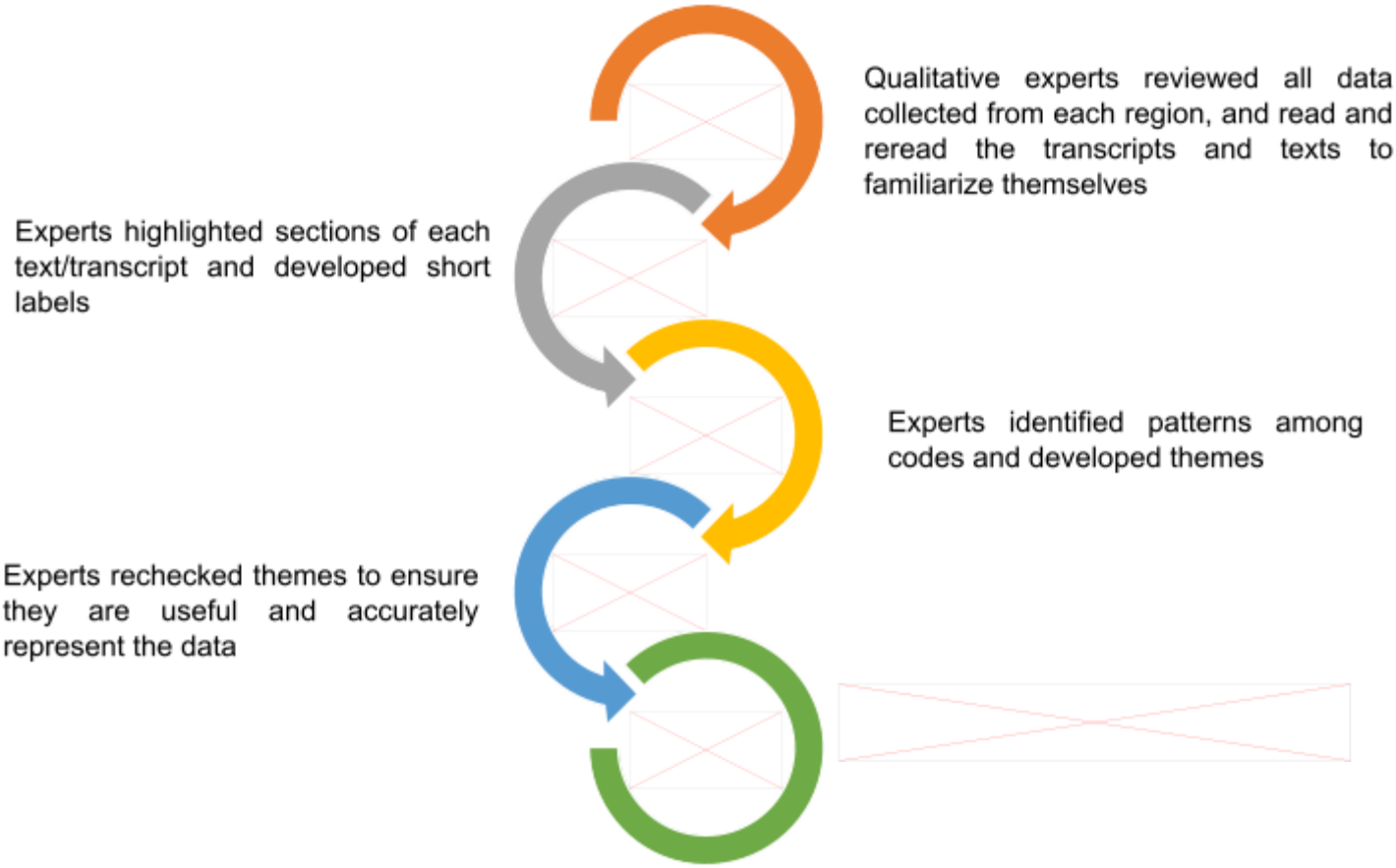


Figure 3. Qualitative data management and analysis processes

4.12 Ethical considerations

Ethical clearance was obtained from the Institutional review board (IRB) of the Ethiopian Public Health Association (EPHA/OG/536/20). RHBs, WoHOs and Kebele administrations were informed about the purpose of the study and permission was secured at all levels before starting field work. All study participants were informed about the purpose, significance, and contents of the study by data collectors and supervisors. Verbal consent was secured from all participants prior to starting the interviews.

5 Findings

The results section is organized into five subsections:

5.1. Coverage and sociodemographic characteristics of participants
5.2. Reaction: Satisfaction of WDA leaders with CBT
5.3. Knowledge and practices of WDA leaders
5.4. Knowledge and practices of WDA network members
5.5. Effect of CBT on HEP-related knowledge and behaviors among WDA leaders and their network members
5.6. The effect of CBT on the functionality of WDA structures
5.7. Facilitators and barriers of CBT

In Section 5.1, we presented the background characteristics of the respondents (WDA leaders and their network members). Section 5.2 deals with information about WDA leaders' satisfaction on different aspects of CBT. In Sections 5.3 and 5.4, we report the findings of the univariate and bivariate analyses of the knowledge and behaviors of WDA leaders and their network members by comparing intervention and comparison groups. In Section 5.5, we report the treatment effect of CBT on knowledge and behaviors of WDA leaders and their network members based on findings from doubly robust estimation techniques involving multivariate analysis. Section 5.6 presents the effect of CBT on the functionality of WDA structures. The last section (5.7) deals with the performance of CBT at woreda level, and barriers to and facilitators of CBT.

5.1 Coverage and sociodemographic characteristics of participants

5.1.1 Study coverage and response rate

We interviewed 187 trained WDA leaders (97.4% response rate [RR]) and 911 network members from intervention areas (97.4% RR), and 374 (97.1% RR) nontrained WDA leaders and 1,825 (95.0% RR) network members from comparison areas.

5.1.2 Background characteristics

Most sociodemographic and economic characteristics of trained and nontrained WDA leaders were similar. However, the mean (SD) age was 36.7 (10.3) for trained WDA leaders, compared to 39.8 (10.1) for nontrained WDA leaders, and the mean (SD) age of husbands was 44.0 (12.4) for trained and 47.2 (12.8) for nontrained WDA leaders. Regarding education, 65% of trained WDA leaders and 45.7% of nontrained WDA leaders had attended school. Community-based health insurance (CBHI) membership status was higher for trained WDA leaders (73.3%) than nontrained WDA leaders (61.8 %; $P < 0.05$). Similarly, compared to network members of nontrained WDA leaders, more women among network members of trained WDA leaders were married (83.5 % vs. 79.5%), CBHI members (58.1 % vs. 50.6 %), younger (mean [SD] age 36.5 [11.2] vs. 38.9 [12.5]), and attended school (45.0 % vs. 38.9%; $P < 0.05$). See Table 3.

Table 3. Background characteristics of WDA leaders stratified by CBT participation, February 2021

Characteristics	WDA Leaders				WDA Members			
	Intervention (N = 187) n (%)	Comparison (N = 374) n (%)	Total (N = 561) n (%)	P-value	Intervention (N = 911) n (%)	Comparison (N = 1,825) n (%)	Total (N = 2736) n (%)	P-value
Region								
Oromia	67 (35.8)	137 (36.6)	204 (36.4)	0.98	337 (37.0)	680 (37.3)	1,017 (37.2)	0.99
Amhara	72 (38.5)	141 (37.7)	213 (38.0)		344 (37.8)	687 (37.6)	1,031 (37.7)	
SNNPR	48 (25.7)	96 (25.7)	144 (25.7)		230 (25.2)	458 (25.1)	688 (25.1)	
Socioeconomic status								
Low	39 (20.9)	90 (24.1)	129 (23.0)	0.34	297 (32.6)	676 (37.0)	973 (35.6)	0.038
Middle	65 (34.8)	142 (38.0)	207 (36.9)		322 (35.3)	635 (34.8)	957 (35.0)	
High	83 (44.4)	142 (38.0)	225 (40.1)		292 (32.1)	514 (28.2)	806 (29.5)	
Marital status								
Married	144 (77.0)	297 (79.4)	441 (78.6)	0.50	761 (83.5)	1,451 (79.5)	2,212 (80.8)	0.004
Widowed	19 (10.2)	41 (11.0)	60 (10.7)		84 (9.2)	215 (11.8)	299 (10.9)	
Other	24 (12.8)	36 (9.6)	60 (10.7)		66 (7.2)	159 (8.7)	225 (8.2)	
Age, mean (SD)	36.7 (10.3)	39.8 (10.1)	35.5 (7.4)	0.003	36.5 (11.2)	38.9 (12.5)	33.2 (7.8)	< 0.001
Husband age, mean (SD)	44.0 (12.4)	47.2 (12.8)	46.1 (12.7)	0.005	44.5 (13.3)	47.2 (14.2)	46.3 (14.0)	< 0.001
Attended school, WDA leaders	93 (65.5)	137 (45.7)	230 (52.0)	< 0.001	342 (45.0)	561 (38.9)	903 (41.0)	0.006
Husband attended school	120 (64.5)	180 (50.4)	300 (55.2)	0.002	449 (50.4)	780 (44.5)	1,229 (46.5)	0.004
Own agricultural land	169 (90.4)	333 (89.0)	502 (89.5)	0.63	678 (88.2)	1,719 (87.4)	2,397 (87.6)	0.70
Agricultural land (in hectares), mean (SD)	0.9 (1.1)	0.9 (0.8)	0.9 (0.9)	0.63	0.8 (1.1)	0.8 (0.8)	0.8 (0.9)	0.30
Household size, mean (SD)	4.9 (1.8)	4.8 (1.9)	4.8 (1.8)	0.37	4.7 (1.9)	4.5 (1.9)	4.6 (1.9)	< 0.001
Livestock, TLU, mean (SD)	4.4 (3.6)	4.2 (3.3)	4.2 (3.4)	0.58	3.6 (3.0)	3.7 (3.2)	3.6 (3.1)	0.16
How long since became WDA leaders (in months), mean (SD)	37.1 (25.6)	41.1 (43.6)		0.25	NA	NA		
HH is currently member of CBHIs	137 (73.3)	231 (61.8)	368 (65.6)	0.007	529 (58.1)	924 (50.6)	1,453 (53.1)	0.001
HH (female)	58 (31.0)	101 (27.0)	159 (28.3)	0.32	218 (23.9)	509 (27.9)	727 (26.6)	0.027
Number of rooms	2.4 (1.1)	2.4 (1.2)	2.4 (1.1)	0.94	2.2 (1.0)	2.2 (1.1)		0.24

Note: *P-values: Pearson chi-square or Fisher exact test; “other” includes single, separated, and divorced; TLUs: tropical livestock units; WDA: Women’s Development Army; CBHI: community-based health insurance; HH: head of household; SD: standard deviation

5.2 Reaction: Satisfaction of WDA leaders with CBT

5.2.1 CBT experiences of trained WDA leaders

Table 4 summarizes self-reported experiences of WDA leaders who received CBT. Although the national CBT guideline suggests the total duration of training should be 52 hours, the mean (SD) self-reported duration of training was 41.9 (14.8). This ranges from 39 hours in Amhara to 46 hours in SNNPR (p = 0.027).

The national CBT guideline specifies that CBT should be concluded by a competency assessment and certification (*i.e.*, after evaluating knowledge and skills related to the seven modules) before graduation. However, this study revealed that only 60.96% of WDA leaders sat for the competency assessment. This was highest in Oromia (79.1%) and lowest in SNNPR (31.3%). This study also revealed that only 23% (of this, only 7.5% were verified by data collectors) received a certificate (Table 4).

Table 4. Duration of training, assessment, and certification status of CBT-trained WDA leaders, February 2021

	Regions				P-value
	Oromia (n = 67)	Amhara (n = 72)	SNNPR (n = 48)	Total (n = 187)	
Reported duration of training (in hours), mean (SD)	42.65 (15.3)	38.58 (15.39)	45.79 (12.03)	41.89 (14.77)	0.027
Received competency assessment, n (%)	53 (79.1)	46 (63.9)	15 (31.3)	114 (60.96)	< 0.001
Status of certification, n (%)					
Certified (verified by data collector)	2 (3.0)	4 (5.6)	8 (16.7)	14 (7.49)	0.025
Certified (not verified by data collector)	6 (8.9)	12 (16.7)	10 (20.8)	28 (14.97)	
Not certified	59 (88.1)	56 (77.8)	30 (62.5)	145 (77.45)	

Note: SD: standard deviation

Qualitative findings reaffirmed that although competency assessment and provision of certificates were deemed important, implementation was not in line with national guidelines. Some WDA leaders said they were evaluated late, or not at all, or not certified formally. A trainer from SNNPR attributed the lack of certification to inadequate implementation at the WoHO level:

> The woreda informed us [trainers] to evaluate them [own trainees] and give them a list of trainees who are efficient. Then I gave them a list of names of competent trainees, but they did not certify them. On our side, we evaluated trainees both in theory and in practice. (HEW, Meskan woreda, Gurage zone, SNNPR)

A key informant WDA leader linked delayed certification to the COVID-19 pandemic:

> The exam was very good. The evaluators were also good. But we did not get a certificate because of the emergency of the COVID-19 pandemic. Additionally, the evaluation and preparation of reassessments for those who failed should be on time. (WDA leader, Mojana-Wodera Woreda, South Wollo zone, Amhara)

At the woreda level, failure to conduct assessments and certifications was reported to be because of lack of adequate budget.

> There was no budget allocated for evaluation and certification; we have not done it. However, the responsibility for preparing the certificate program was given to the Health Extension workers. But we have not followed up with the HEWs whether they have accomplished this task or not. Hence, I am sure that there were gaps in providing the certificate. This is our major gap. (HEP coordinator, Bati woreda, Oromia special zone, Amhara)

The absence of evaluation and certification resulted in dissatisfaction and eroded trained WDA leaders' trust in healthcare sector staff.

> Lack of evaluation and certification created a negative impression. Initially, we gave a brief orientation about training time, the content, evaluation, CoC processes, and certification. They eagerly waited for the certificate. However, trainees did not sit for formal assessments and never got certified. Now, when we communicate with them for another job, they complain that we did not keep our promise. (HEP coordinator, Sinana woreda, Bale zone, Oromia)

When trained WDA leaders reported having passed competency assessments, they reported different forms of assessments, highlighting a lack of consistency across regions, woredas, and kebeles. While some reported that they completed individual oral or written assessments, others reported that they were evaluated as a group. One of the key informants said,

> Actually, the evaluation was done in a group, which might not show who responded well and who did not. I think the score was given as a group because there were participants who did not answer but still passed the exam. The examiner had little knowledge about who is doing well within a group. But I still recommend it to be in a group because mothers might be afraid of being interviewed alone, and they would be confident when the examination is given in a group. However, understanding individual knowledge is also important. (WDA leader, Enamor Ener woreda, Gurage zone, SNNPR)

5.2.2 WDA leaders' satisfaction with CBT

Almost all participants (98.6%) considered CBT useful. Responding to a question on how CBT was useful for them, a large majority reported that CBT improved their knowledge, ability to disseminate health messages, communication skills, and healthy living practices (Figure 4).

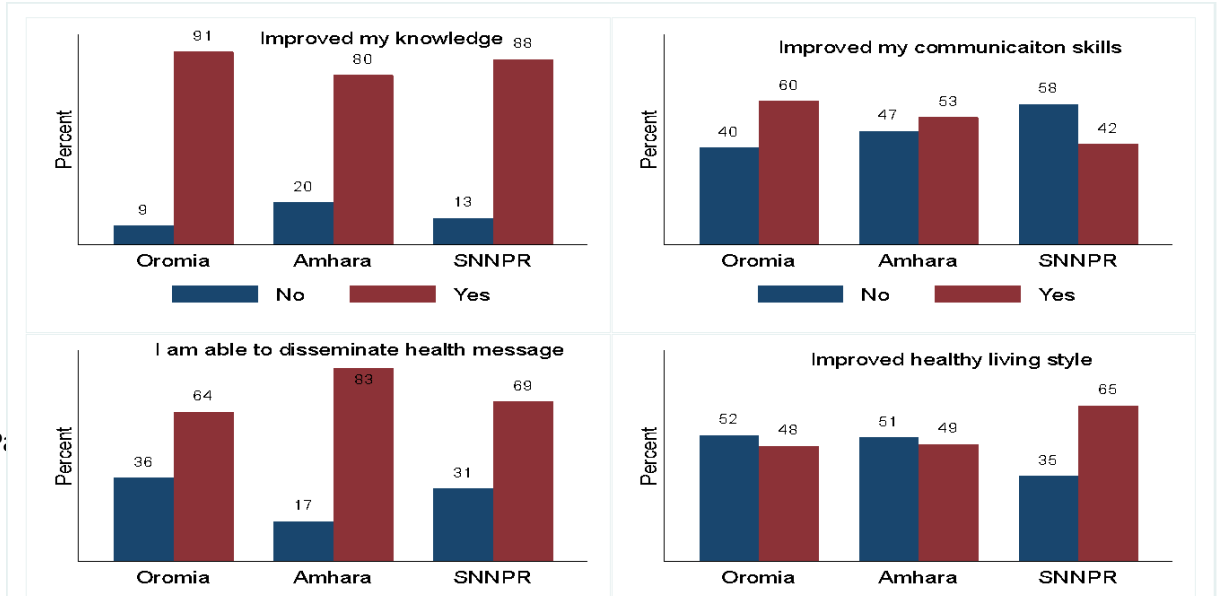


Figure 4. WDA leaders' reactions regarding the usefulness of CBT, February 2021

The mean (SD) satisfaction score was 3.13 (0.39) on a 4-point scale. Overall, 62 (33.1%) of trained WDA leaders were highly satisfied with the training. In all other domains of satisfaction, more than 50% of participants declared that the training was excellent (Figure 5). Figure 5 shows between and within region differences in satisfaction score. Accordingly, although WDA leaders from SNNPR reported greater satisfaction, there is also higher variability within the region. In contrast, WDA leaders from Oromia rated satisfaction lower than those in the other regions, and there is also substantial variability within the region.

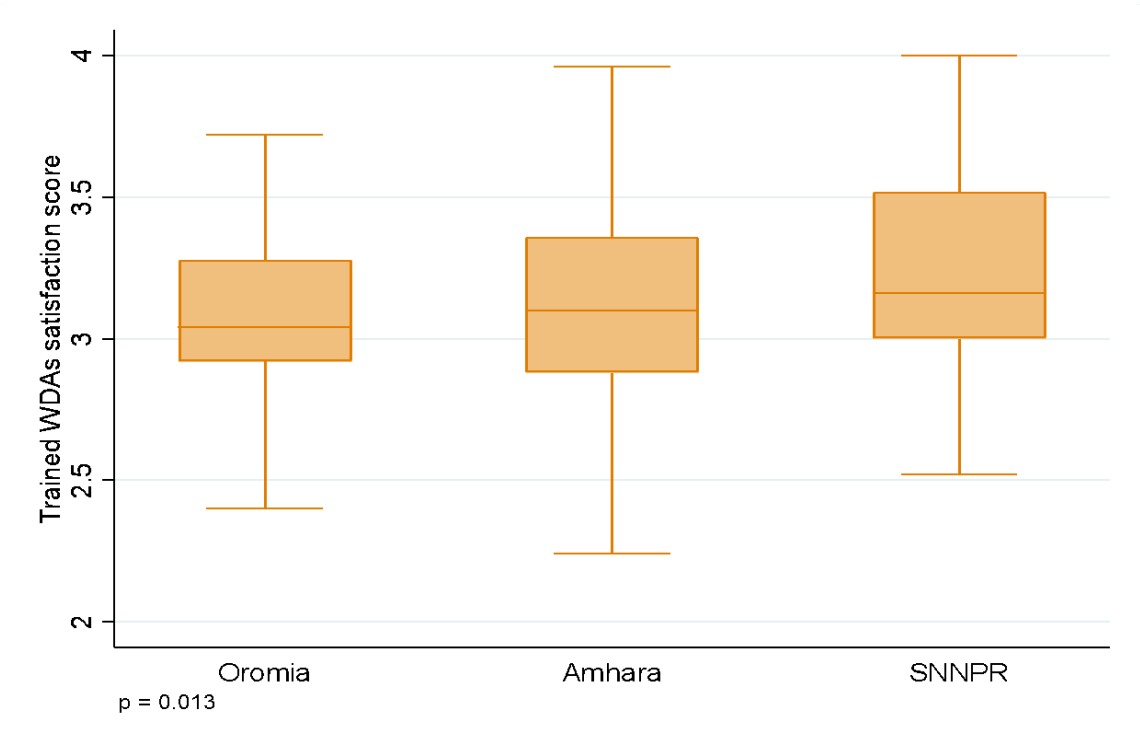


Figure 5. Satisfaction of trained WDA leaders with CBT program, February 2021

Average satisfaction of WDA leaders with CBT showed minor difference across age groups. However, variability in level of satisfaction was highest among the youngest WDA leaders (Figure 6).

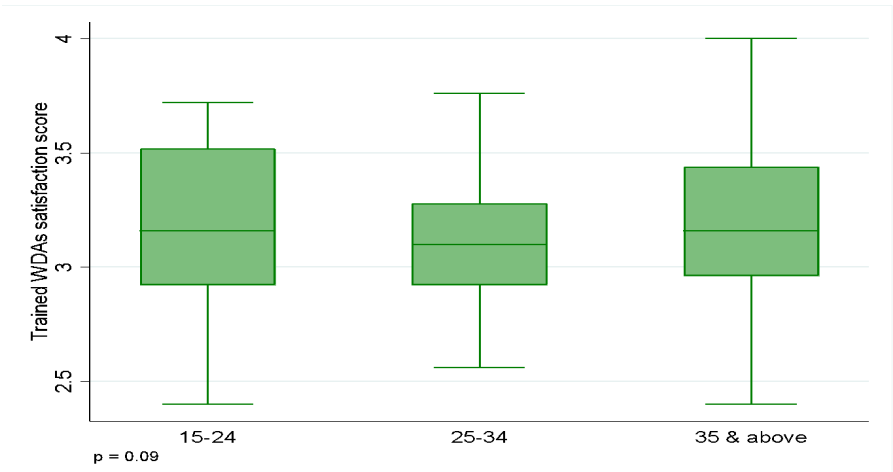


Figure 6 shows between and within age differences in satisfaction scores of WDA leaders. Younger WDA leaders rated satisfaction higher, although it is not statistically significant. There is also higher variability within this age group. WDA leaders aged 15–24 rated this lower than other groups.

Figure 6. WDA leaders' satisfaction with CBT program by age group, February 2021

Findings from in-depth interviews also showed that WDA leaders were satisfied with the process, contents, and mode of delivery. Most trained WDA leaders have a positive attitude towards the CBT program. One of the Key Informants said:

> The training was satisfying. The HEWs trained us based on the modules and they gave us satisfactory answers whenever we asked questions. I like the training because it helps mothers reduce risk of death from hemorrhage. (WDA leader, Basona woreda, North Shewa zone, Amhara)

Because of the value WDA leaders placed on CBT and their level of satisfaction, WDA leaders were committed to attending CBT despite challenges:

> Although women have other competing priorities and workload related to household chores, they were committed to attend the training because they were happy with the training. (WDA leader, Meskan woreda, Gurage zone, SNNPR)

Trained WDA leaders were also asked which parts of the training process they liked. Most favorably rated training content, mode of delivery, and trainers, but timing of the training and adequacy of training materials were reported as likeable less frequently (Figure 7).

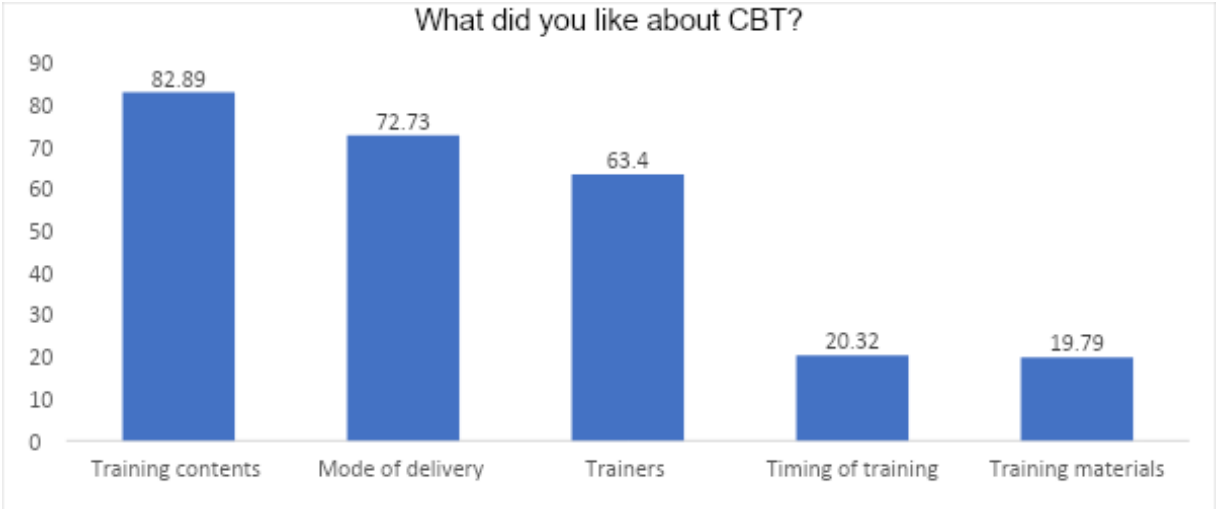


Figure 7. WDA leaders' reactions to CBT processes, February 2021

In contrast, respondents rated issues related to adequacy of the training duration, suitability of training days considering the daily routines of WDA leaders, and the adequacy of training materials as relatively unsatisfactory (Table 5).

Table 5. Summary statistics of CBT-trained WDA leaders' reactions to CBT, February 2021

Items	Oromia (N = 67) Mean (SD)	Amhara (N = 72) Mean (SD)	SNNPR (N = 48) Mean (SD)	Total (n = 187)
Training included topics that can help us solve our knowledge gaps as WDA leaders	3.13 (.54)	3.06 (.67)	3.47 (.54)	3.19 (.62)
Training included topics that can help us solve our skill gaps as WDA leaders	3.16 (.29)	3.09 (.75)	3.25 (.44)	3.16 (.66)
Training included topics that can help us solve our attitude gaps as WDA leaders	3.07 (.58)	3.13 (.69)	3.25 (.63)	3.14 (.64)
Training could enable us to solve the knowledge gaps of our network members	3.04 (.58)	3.16 (.67)	3.20 (.76)	3.13 (.70)
Training could enable us to solve the skills gaps of our network members	2.97 (.69)	3.13 (.65)	3.02 (.75)	3.04 (.69)
Training could enable us to solve the attitude gaps of our network members	2.80 (.65)	3.22 (.61)	3.22 (.69)	3.07 (.67)
Training days suited to our routine life activities	3.02 (.67)	2.70 (.97)	2.68 (1.07)	2.81 (.91)
Time allocated for the training was enough	2.91 (.84)	2.59 (1.04)	3.00 (.89)	2.81 (.95)
Training venue was suitable	3.01 (.76)	3.06 (.69)	3.10 (.88)	3.05 (.77)
Resources and training materials were important and sufficient	2.76 (.79)	2.68 (.94)	2.70 (1.20)	2.71 (.96)
Delivery modality of training was appropriate	3.11 (.61)	3.08 (.59)	3.20 (.58)	3.12 (.59)
Trainers had good expertise/knowledge of the subject matter	3.04 (.56)	3.09 (.69)	3.18 (.86)	3.10 (.69)
Trainers had good delivery and communication skills	3.13 (.60)	3.22 (.61)	3.22 (.83)	3.19 (.66)
Training was delivered by accounting cultural and social conditions	3.05 (.71)	3.12 (.71)	3.31 (.62)	3.14 (.69)

Training has improved my knowledge about HEP packages	3.19 (.55)	3.27 (.61)	3.60 (.49)	3.33 (.58)
Training has improved my attitude towards on HEP packages	3.14 (.52)	3.20 (.57)	3.50 (.50)	3.24 (.61)
Training has improved my attitude towards WDA network	3.08 (.51)	3.23 (.51)	3.31 (.68)	3.26 (.55)
Training has improved my health behavior	3.14 (.55)	3.27 (.50)	3.43 (.50)	3.20 (.56)
After the training, I am determined to positively change the health behavior of my network members	3.14 (.65)	3.15 (.59)	3.37 (.48)	3.20 (.59)
Training has improved my motivation to work as a WDA leader	3.10 (.76)	3.08 (.70)	3.31 (.71)	3.27 (.53)
Training has improved my self-efficacy to work as a WDA leader	3.13 (.60)	3.18 (.67)	3.25 (.75)	3.06 (.68)
Training has improved my communication skills	3.11 (.47)	3.23 (.66)	3.31 (.62)	3.14 (.67)
I recommend this training for others	3.38 (.67)	3.23 (.68)	3.54 (.65)	3.36 (.67)
Overall satisfaction score, mean SD	3.07 (.33)	3.10 (.43)	3.26 (.39)	3.12 (.39)

Key informant interviews revealed that informants' satisfaction stemmed from *the knowledge and skills* or self-efficacy they developed from the training. As one of the WDA leaders said,

> I am very satisfied with the knowledge I got and that I can educate others who do not have the same knowledge, and to my children as well. For example, we were taught about child feeding practices and how to properly prepare child foods using food materials we have at home. (WDA leader, Meskan woreda, Gurage zone, SNNPR)

Trained WDA leaders were also satisfied with CBT because of trainers' approaches and improvements in their relationships with trainers.

> I am happy with the method, the manual, content, and approach including practical-based training. I am telling you frankly the approach of the trainer was the most attractive part. She was friendly. We were following the training with high interest and attention. We engaged actively and always look forward to the next training dates. (WDA leader, Tiyo woreda, Arsi zone, Oromia)

In terms of training content, we observed that although it differs across regions, topics on WASH, ANC, and FP were the most liked (Figure 8).

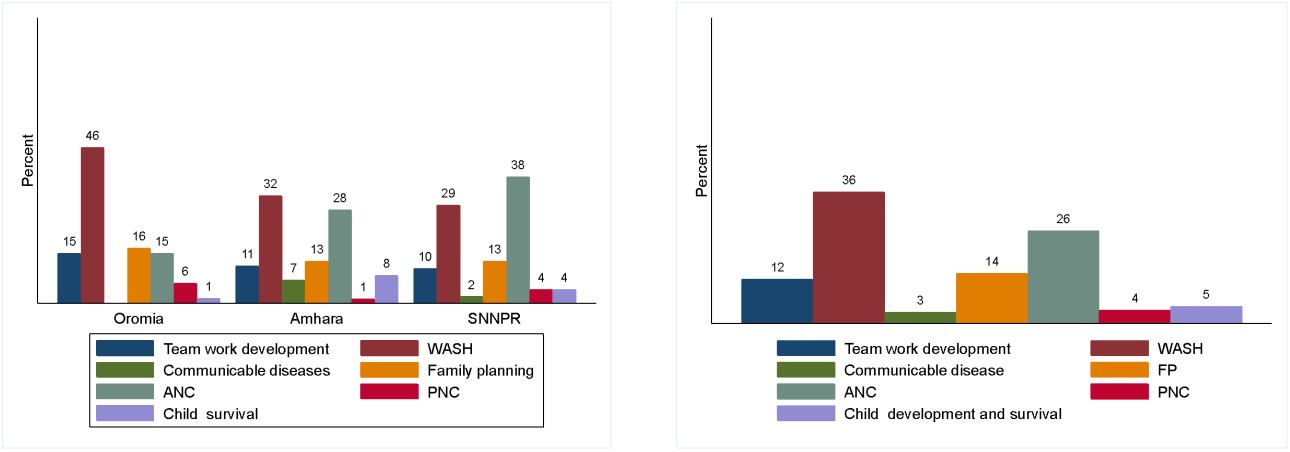


Figure 8. WDA leaders' reactions to CBT module contents, February 2021

Practical or demonstration sessions were mostly viewed favorably. One of WDA leaders described this by saying,

> The training had practical session like how to prepare child food from a mix of vegetables and cereals, and demonstrations on the different types of contraceptives such as injectables, IUCD, pills. So, I am happy with the teaching method. (WDA leader, Tiyo woreda, Arsi zone, Oromia)

In areas where the training materials were adequate and complete, WDA leaders acknowledged that the training manuals (family health guides) were good in terms of color, font, and figures. The pictures and figures included in the manual are self-explanatory, especially for those trainees who cannot read and write. They also appreciated the importance of using local languages:

> The pictures are understandable. For example, pictures that showed how respiratory diseases are transmitted from one person to another by drawing the virus spreading in the air when a patient cough. Additionally, the training modules were prepared in Afan Oromo; the fonts are bold and visible. These manuals were highly attractive for us. (WDA leader, Sinana woreda, Bale zone, Oromia)

A WDA leader from Amhara also mentioned the use of color-printed manuals as having encouraged them to read and observe the contents:

> If there is a picture in the manual that is printed in a unique color, we like to focus on the colored part and ignore other parts of the manual. However, if it has different colors like our manual, we could see all parts of the material. (WDA leader, Basona woreda, North Showa zone, Amhara)

However, these positive features of the training manuals were not available uniformly. Some key informant interview participants complained about a shortage of color-printed manuals and missing pages in their versions.

Training venue and timing were the other aspects of training that emerged as factors determining the satisfaction of WDA leaders. Key informants, both trainers and trainees, mentioned that training date and time should be arranged based on the trainees' suggestions. WDA leaders took the training for two or three days per week.

Convenience of venue and timing of the training was not consistent across settings. In some places, flexibility and participation of trainees in determining venues and training periods was appreciated as a strength of the program, while in other places, overlaps between the training period and harvesting seasons and inconvenience of training venues were sources of dissatisfaction.

> We [trainees] decided the time and date of the training so that it did not affect our work, and the training should not also create pressure on pregnant mothers. As per our agreement, we attended the training two days a week. (WDA leader, Shirka woreda, Arsi Zone, Oromia)

> The training was arranged for two hours per day in the morning, and it should not have been discontinued. However, if they had other responsibility at that time, we discussed the issue and arranged the training for the afternoon. The time flexibility helps them very much because it prevents the risk of overlap with other activities. (HEW, Meskan woreda, Gurage zone, SNNPR)

Another trainee from Oromia mentioned the importance of considering the season:

> In our case, the training was given during the harvesting time, which hindered training participation. During that time, women were busy harvesting or collecting and cutting peas, beans, barley, and wheat. The women loved to learn even from dawn to dusk, yet the time was not appropriate to do that, so we were forced to interrupt the training. (WDA leader, Sinana woreda, Bale Zone, Oromia)

5.3 Knowledge and practices of WDA leaders

We assessed the knowledge and behavior of trained and nontrained WDA leaders using indicators related to WASH, DPC, and family health.

5.3.1 Water, sanitation, and hygiene

Although not all indicators showed significant differences, WDA leaders who received CBT had more knowledge about WASH than nontrained WDA leaders. Trained WDA leaders were more knowledgeable about the importance of using toilets, the criteria of a healthy home, and methods used to prevent insect and rodent infestation than did nontrained WDA leaders ($p < 0.05$; Table 6).

Table 6. WASH facilities and knowledge status among CBT-trained and nontrained WDA leaders, February 2021

Knowledge areas related to WASH	Trained WDA leaders, N = 187, n (%)	Nontrained WDA leaders, N = 374, n (%)	Total (N = 561) n (%)	P-value
The importance of using toilets Prevent communicable diseases	165 (88.2)	313 (83.7)	478 (85.2)	0.15
Prevent water and food poisoning	91 (48.7)	153 (40.9)	244 (43.5)	0.08
Protect the environment from bad odors and pathogens	146 (78.1)	271 (72.5)	417 (74.3)	0.15
Criteria for a healthy home				
Free from pathogens	139 (74.3)	236 (63.1)	375 (66.8)	0.008
Does not cause physical harm	58 (31.0)	99 (26.5)	157 (28.0)	0.26
Reduces social and psychological stress	41 (21.9)	76 (20.3)	117 (20.9)	0.66
Has shelf for food and water utensils	103 (55.1)	167 (44.7)	270 (48.1)	0.020
The house should be far from toilets/separate/kitchen or pets	139 (74.3)	253 (67.6)	392 (69.9)	0.10
Methods of preventing insect and or rodent infestations				
Keep the environment clean and dry and use sewage disposal	131 (70.1)	210 (56.1)	341 (60.8)	0.001
Eliminate insect breeding areas	72 (38.5)	114 (30.5)	186 (33.2)	0.057
Eliminate predatory animals	122 (65.2)	243 (65.0)	365 (65.1)	0.95
Keep food away from such places	61 (32.6)	110 (29.4)	171 (30.5)	0.44
Adequate quality and safe water supply	27 (14.4)	43 (11.5)	70 (12.5)	0.32

Note: *P-values: Pearson's chi-squared or Fisher's exact test

Like differences in WASH-related knowledge, compared to nontrained WDA leaders, trained WDA leaders had better WASH related practices. The use of a toilet, participating in WASH campaigns, water treatment (*i.e.*, use of chlorine or bleach to treat water), having a functional solid waste disposal (SWD) pit, and the use of a liquid waste disposal system were higher among trained than nontrained WDA leaders ($p \leq 0.05$). However, no statistically significant difference was observed between the two groups

in relation to a number of other WASH-related practices targeted by CBT, including having a separate kitchen, separating animals from living rooms, and handwashing practices during critical times (Table 7).

Table 7. Practices related to WASH among CBT-trained and nontrained WDA leaders, February 2021

WASH-related practices	Trained WDA leaders, N = 187, n (%)	Nontrained WDA leaders, N = 374, n (%)	Total (N = 561), n (%)	P-value
Handwashing practices (critical times)				
Before preparing foods	171 (91.4)	321 (85.8)	492 (87.7)	0.056
While preparing foods	132 (70.6)	250 (66.8)	382 (68.1)	0.37
After using toilet	147 (78.6)	272 (72.7)	419 (74.7)	0.13
After cleaning child who used the toilet	70 (37.4)	101 (27.0)	171 (30.5)	0.011
After blowing nose, coughing, or sneezing	34 (18.2)	53 (14.2)	87 (15.5)	0.22
After touching an animal, animal feed, or animal waste	75 (40.1)	148 (39.6)	223 (39.8)	0.90
After handling animal food or treatments	52 (27.8)	123 (32.9)	175 (31.2)	0.22
Toilet use				
Open defecation	31 (16.6)	88 (23.5)	119 (21.2)	0.057
Pit latrine without slab	134 (71.7)	252 (67.4)	386 (68.8)	
Pit latrine with slab	20 (10.7)	34 (9.1)	54 (9.6)	
Other	2 (1.1)	0 (0.0)	2(0.4)	
Improved sanitation facilities	22 (14.1)	34 (11.9)	56 (12.7)	0.50
Access to water and water treatment				
Treat water	41 (22.0)	41 (11.1)	82 (14.8)	<0.001
Add chlorine or bleach	25 (61)	16 (39)	41 (50)	0.047
Time to fetch water, in minutes, m (SD)	29.4 (21.9)	33.9 (33.8)	32.4 (30.5)	0.15
Improved sources of drinking water	169 (90.4)	306 (81.8)	475 (84.7)	0.008
Household WASH facilities				
Water is available for handwashing in the compound	72 (38.5)	113 (30.2)	185 (33.0)	0.049
HH has shelf on which feeding utensils can be stored	144 (77.0)	288 (77.0)	432 (77.0)	1
Feeding utensils covered and protected from flies	144 (77.0)	273 (73.0)	417 (74.3)	0.31
Have a separate kitchen room	153 (81.8)	307 (82.1)	460 (82.0)	0.94
Have a separate sleeping room	126 (67.4)	263 (70.3)	389 (69.3)	0.48
Have a separate room for livestock	143 (76.5)	286 (76.5)	429 (76.5)	1
HH participated in WASH campaign	56 (29.9)	64 (17.1)	120 (21.4)	0.001

WASH-related practices	Trained WDA leaders, N = 187, n (%)	Nontrained WDA leaders, N = 374, n (%)	Total (N = 561), n (%)	P-value
HH has functional solid waste disposal (SWD)	89 (47.6)	119 (31.8)	208 (37.1)	0.001
HH has functional liquid disposal system	49 (26.2)	70 (18.7)	119 (21.2)	0.041

Note: *P-values: Pearson chi-squared or Fisher's exact test; HH: Household; SD: Standard Deviation;

5.3.2 Disease prevention and control

WDA leaders who had received CBT had better knowledge and perceptions about TB, HIV/AIDS, NCDs, and cervical cancer than nontrained WDA leaders. For instance, 89.9% of trained WDA leaders knew that TB can be cured, while 83.1% of nontrained WDA leaders did (p = 0.039). Similarly, knowledge of condom uses to reduce the chance of getting HIV was 84.9% among trained WDA leaders, compared to 74.7% among nontrained leaders (p = 0.006). These differences were observed for knowledge of prevention of mother-to-child transmission, information about cervical cancer, and knowledge of cervical cancer prevention (Table 8 and Table 9).

Table 8. Knowledge about TB and HIV/AIDS among CBT-trained and nontrained WDA leaders, February 2021

Characteristics	Trained WDA leaders, (N = 187), n (%)	Nontrained WDA leaders, (N = 374) n (%)	Total (N = 561) n (%)	P-value
Knowledge and perceptions about TB				
Heard about TB	178 (95.2)	332 (88.8)	510 (90.9)	0.013
Know TB can be cured	160 (89.9)	276 (83.1)	436 (85.5)	0.039
Knowledge and perceptions about HIV/AIDS				
Heard of HIV/AIDS	185 (98.9)	359 (96.0)	544 (97.0)	0.055
Believe people can reduce chance of getting HIV by having just one uninfected sex partner who has no other sex partners	163 (88.1)	285 (79.4)	448 (82.4)	0.011
Believe people can get HIV from mosquito bites	65 (35.1)	132 (36.8)	197 (36.2)	0.71
Believe people can reduce their chance of getting HIV by using a condom every time they have sex	157 (84.9)	268 (74.7)	425 (78.1)	0.006
Believe people can get HIV by sharing food with a person who has HIV	169 (91.4)	303 (84.4)	472 (86.8)	0.023
Believe it possible for a healthy-looking person to have HIV	160 (86.5)	286 (79.7)	446 (82.0)	0.05
Heard of ARVs to treat HIV	143 (77.3)	260 (72.4)	403 (74.1)	0.22
Know about special medicines a doctor or a nurse can give a woman infected with HIV to reduce the risk of transmission to the baby	162 (87.6)	277 (77.2)	439 (80.7)	0.004

Note: *P-values: Pearson chi-squared or Fisher exact test

Table 9. Knowledge about NCDs, cervical cancer, adulthood illness, and treatment-seeking behaviors among CBT-trained and nontrained WDA leaders, February 2021

Characteristics	Trained WDA leaders, n (%)	Non- trained WDA leaders, n (%)	Total, n (%)	P-value
Knowledge and perceptions about NCDs				
Heard of NCDs such as hypertension, DM, cancer	160 (85.6)	282 (75.4)	442 (78.8)	0.006
Heard of cervical cancer	145 (77.5)	200 (53.5)	345 (61.5)	<0.001
Source of information about NCDs				
Health Extension workers	113 (70.6)	98 (34.8)	211 (47.7)	<0.001
WDA leaders	3 (1.9)	14 (5.0)	17 (3.8)	
Other health professionals from health centers or hospitals	14 (8.8)	55 (19.5)	69 (15.6)	
Media	18 (11.3)	67 (23.8)	85 (19.2)	
Other	12 (7.5)	48 (17.0)	60 (13.6)	
Prevention and control methods of cervical cancer				
Avoid unprotected sexual intercourse	42 (29.0)	38 (19.0)	80 (23.2)	0.030
Vaccination	71 (49.0)	67 (33.5)	138 (40.0)	0.004
Screening	92 (63.4)	127 (63.5)	219 (63.5)	0.99
Other	12 (8.3)	30 (15.0)	42 (12.2)	0.059
Treatment-seeking behaviors (household members)				
At least one sick household member in the last 12 months	50 (26.7)	115 (30.7)	165 (29.4)	0.33
Treatment sought for sick member (most recent)	46 (92.0)	98 (85.2)	144 (87.3)	0.23
Place where treatment was sought				
Health center or health posts	24 (52)	54 (55)	78 (54.2)	0.61
Hospital	15 (33)	24 (24)	39 (27.1)	
Private clinics	6 (13)	19 (19)	25 (17.4)	
Other	1 (2)	1 (1)	2(1.4)	

Note: *P-values: Pearson chi-squared or Fisher exact test; NCD: Non-Communicable Diseases; DM; Diabetes Mellitus

5.3.3 Family Health

Family planning

Knowledge of FP methods was in general high for injectables, implants, and combined contraceptive pills among both trained and nontrained WDA leaders. For the other methods of contraceptives, with which the general population is relatively less familiar, knowledge about methods was higher among trained

WDA leaders than nontrained WDA leaders. The difference was more pronounced for methods that are less familiar to the general population (Figure 9).

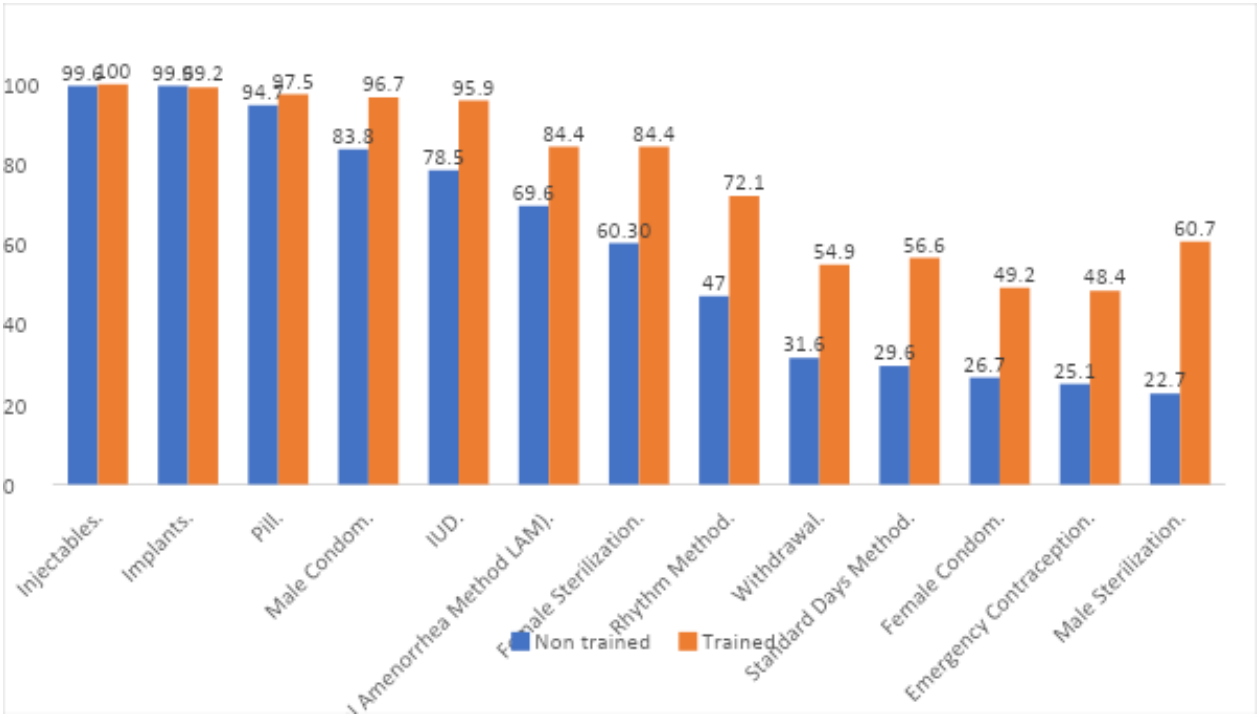


Figure 9. Knowledge of FP methods in trained and nontrained WDA leaders, February 2021

Having ever used FP methods was reported by 45.0% of trained WDA leaders and 47.8% of nontrained WDA leaders of reproductive age. The prevalence of modern contraceptive (mCPR) use was higher among nontrained WDA leaders (44.1%) than among trained WDA leaders (39.3%). The proportion of WDA leaders who used LARC was almost similar among trained (32.0%) and nontrained (31.8%) WDA leaders (Table 10).

Method mix was relatively better among trained WDA leaders than nontrained ones (Figure 10). In relative terms, LARC was more commonly used in Amhara (37% among nontrained and 67% among trained WDA leaders) and SNNPR (42% among nontrained and 23% among trained WDA leaders) than in Oromia (22% among nontrained and 20% among trained WDA leaders).

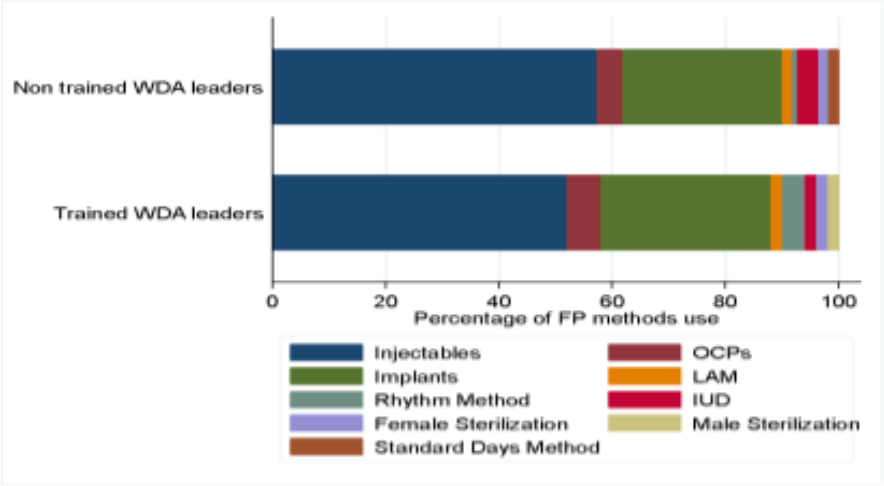


*Figure SEQ Figure * ARABIC 10. Contraceptive method mix among current users of FP among trained and nontrained WDA leaders*

Table 10. FP practices among WDA leaders by training status, February 2021

Characteristics	Trained WDA leaders, n (%)	Nontrained WDA leaders, n (%)	Total n (%)	P-value
Reproductive history				
Ever been pregnant	120 (98.4)	238 (96.4)	358 (97.0)	0.29
Currently pregnant	11 (9.2)	17 (7.1)	28 (7.8)	0.62
Current pregnancy (wanted)	9 (82)	17 (100)	26 (93)	0.068
Most recent pregnancy wanted	102 (85.0)	212 (89.1)	314 (87.7)	0.16
No, wanted later	10 (8.3)	20 (8.4)	30 (8.4)	
Not wanted at all	8 (6.7)	6 (2.5)	14 (3.9)	
Main sources of information about FP				
Health Extension workers	101 (82.2)	156 (64.2)	257 (70.4)	<0.001
Health professionals	18 (14.8)	40 (16.5)	58 (15.9)	
Other sources	3 (2.5)	47 (19.3)	50 (13.7)	
Practices related to FP				
Ever use FP (any FP methods)	50 (45.0)	110 (47.8)	160 (46.9)	0.63
Current use, modern CPR	48 (39.3)	109 (44.1)	157 (42.5)	0.38
Current use, LARC	16(32.0)	35 (31.8)	51 (31.9)	0.98
How long since used the current FP methods (in months)	28.6 (41.9)	27.5(31.5)	27.8 (35.0)	0.86
Used (anything or tried) any FP method in any way to delay or avoid getting pregnant in the last 12 months	57 (46.7)	109 (44.1)	166 (45.0)	0.64
Source of FP				
Health center	19(38.0)	59 (53.6)	78 (48.8)	0.059
Health post	22 (44.0)	43 (39.1)	65 (40.6)	
Other	9 (18.0)	8(7.3)	17 (10.6)	

Note: *P-values: Pearson chi-squared or Fisher exact test; FP: Family Planning

Maternal health

WDA leaders were asked about danger signs during pregnancy. Most respondents mentioned vaginal bleeding, gushing fluids, and lower abdominal pain as danger signs in pregnancy. Fever and convulsions were mentioned less frequently (Figure 11).

Knowledge of most of the danger signs in pregnancy was higher for trained WDA leaders than their nontrained counterparts (83% vs. 74% for vaginal bleeding, 45% vs. 40% for gushing vaginal fluids, 45% vs. 25% for edema, 56% vs. 31% for severe headache, 40% vs. 21% for fever, and 33% vs. 11% for convulsions), except for lower abdominal pain, which was mentioned more frequently among nontrained WDA leaders (38%) than among trained ones (34%; Figure 12).

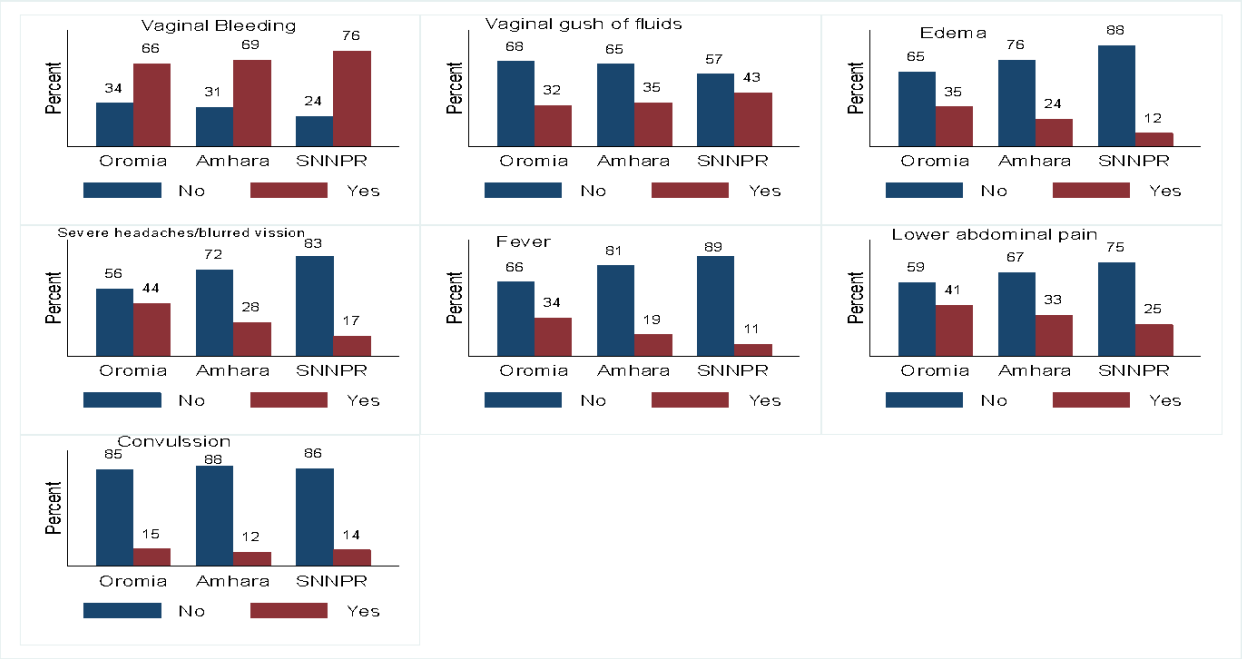


Figure 11. Awareness about danger signs during pregnancy among trained and nontrained WDA leaders

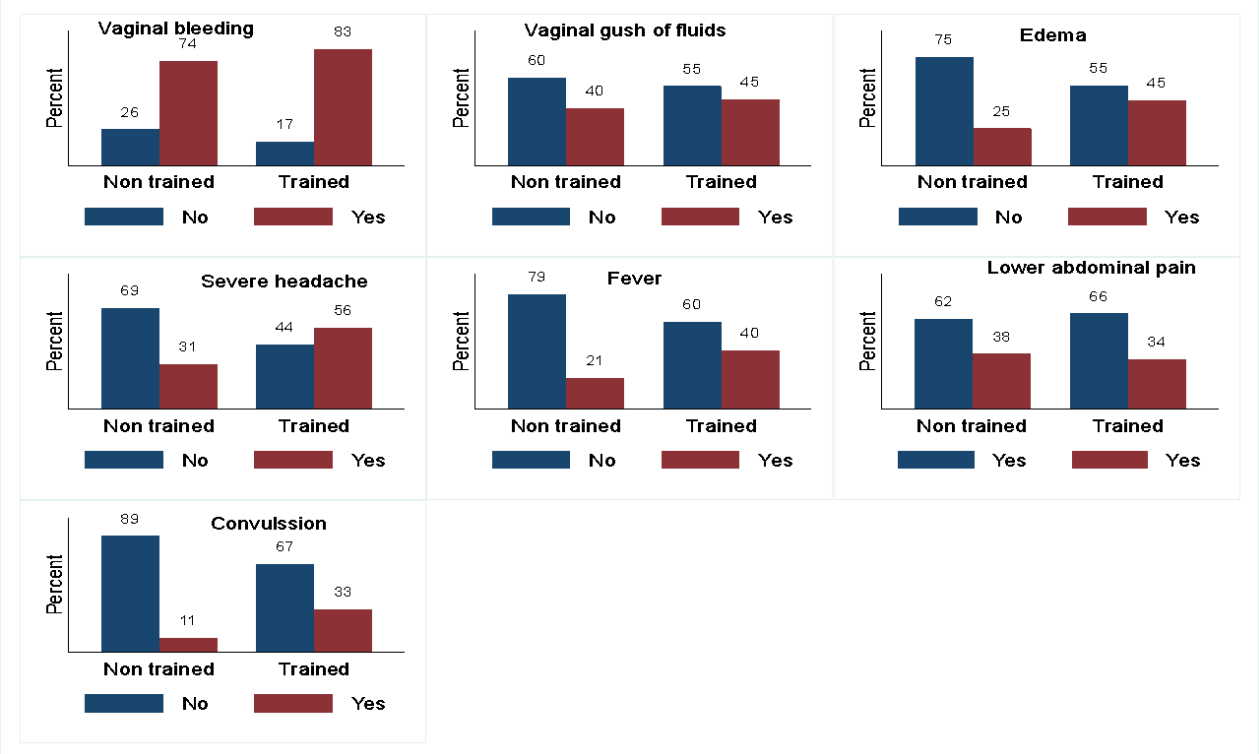


Figure 12. Awareness about danger signs during pregnancy among trained and nontrained WDA leaders by CBT status

We also asked WDA leaders about the recommended time to have the first PNC visit. According to 35% of trained WDA leaders and 32% of nontrained WDA leaders, the first PNC visit should be provided within 24 hours (Figure 13).

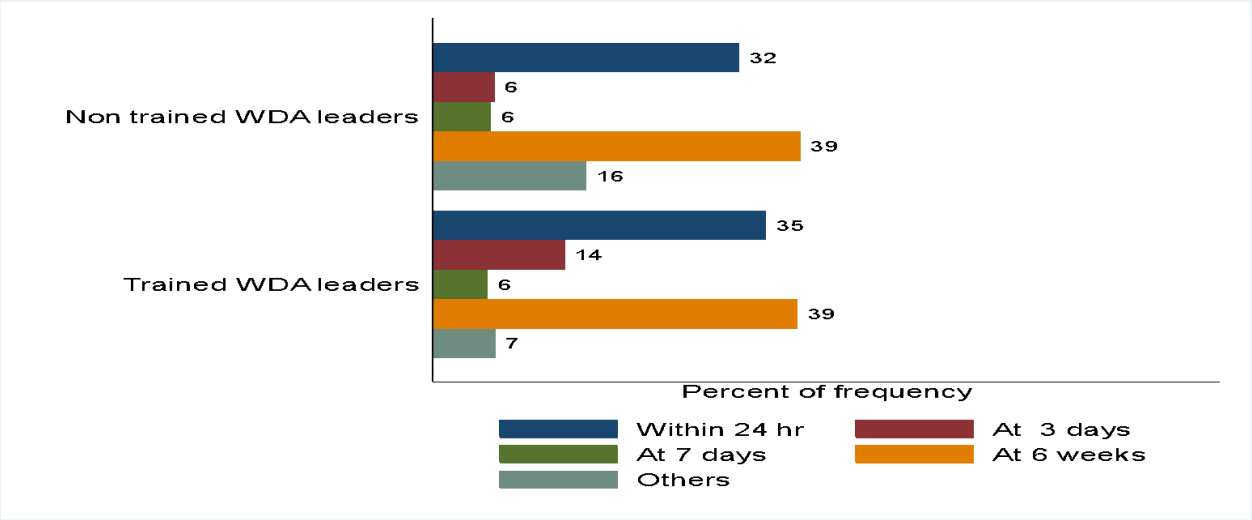


Figure 13. WDA leaders' knowledge about timing of first PNC visit, February 2021

Table 11 presents the findings on ANC, delivery, and PNC among WDA leaders. Ninety-one percent of trained WDA leaders and 86% of nontrained WDA leaders received ANC during their most recent pregnancy, and 72% of the trained WDA leaders and 58% of nontrained WDA leaders had at least four ANC visits during their last pregnancy. We also found that 55.0% of trained WDA leaders, compared to 43.3% of nontrained WDA leaders, delivered their most recent children at healthcare facilities (p = 0.036). Postnatal checkups were more common among trained than nontrained WDA leaders (Table 11).

Table 11. Knowledge and practices related to maternal and child healthcare services among WDA leaders, February 2021

Characteristics	Trained WDA leaders, n (%)	Nontrained WDA leaders, n (%)	Total n (%)	P-value
Maternal health care services: practices				
Received at least one ANC for most recent pregnancy	29 (91)	24 (86)	53 (88)	0.55
Received 4+ ANC visits for most recent pregnancy	21 (72)	14 (58)	35 (66)	0.28
Place ANC services received from				
Health center	15 (52)	14 (58)	29 (55)	0.58
Health post	12 (41)	7 (29)	19 (36)	
Hospital	2 (7)	3 (13)	5 (9)	
Place of birth for most recent delivery				
Home	54 (45.0)	135 (56.7)	189 (52.8)	0.036
Health facility	66 (55.0)	103 (43.3)	169 (47.2)	
Received PNC for most recent delivery	23 (72)	15 (54)	38 (63)	0.14

Note: *P-values: Pearson chi-squared or Fisher exact test

Child vaccination

Respondents had 134 children in the age range of 12–23 months. Among these, 76 (56.7%) were reported to have a vaccination card, booklet, or other home-based record. Of these, the interviewers could verify this information for 55 (72.0%). A significant variation appeared in having a vaccination card between trained and nontrained WDA leaders: 63.2% (36/57) among trained WDA leaders and 52.0% (40/77) among nontrained WDA leaders. Based on information from vaccination cards, 53% of nontrained WDA leaders' children and 43% of trained WDA leaders' children aged 12–23 months received all basic vaccinations at some time. Although not statistically significant, the difference in vaccination coverage between children of trained and nontrained WDA leaders is highest for the first dose of BCG and the polio, pentavalent, and rotavirus 2 vaccines (Figure 14).

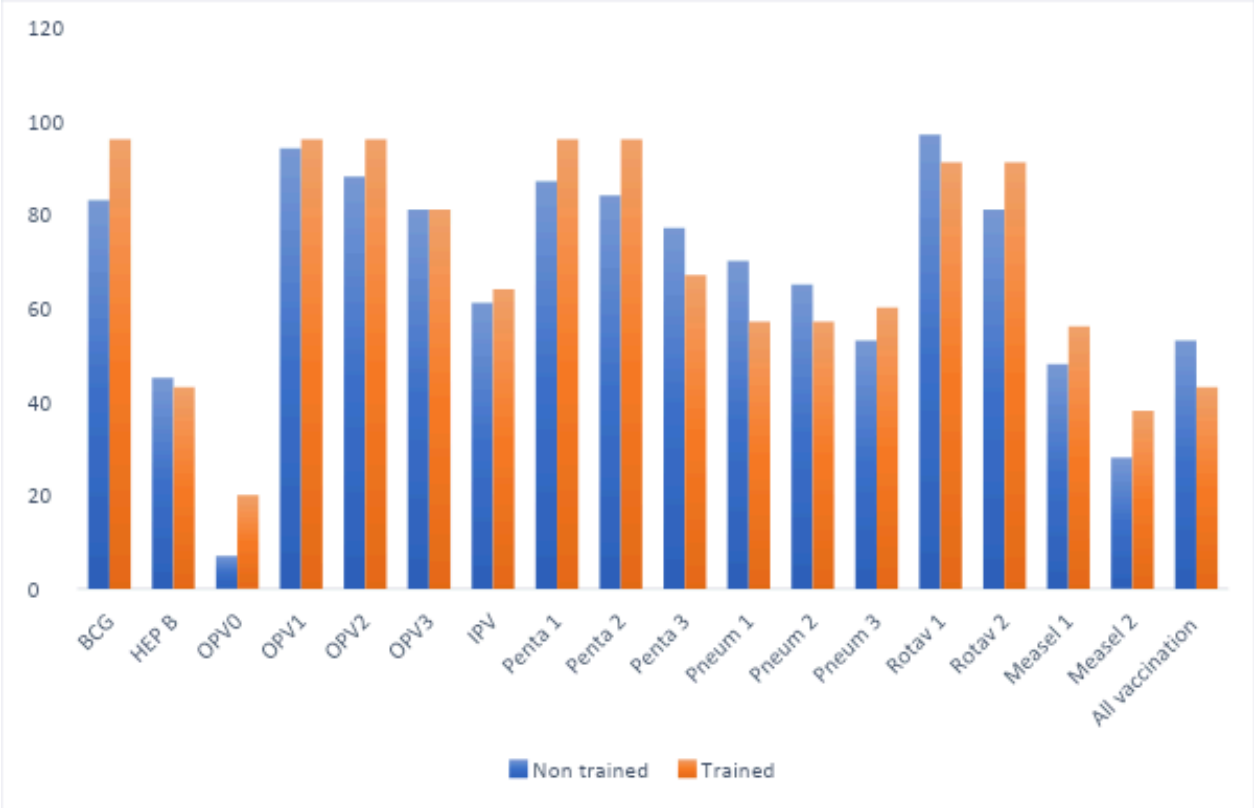


Figure 14. Vaccination status among children of trained and nontrained WDA leaders (12–23 months age), February 2021

Treatment seeking for childhood illnesses

During the two weeks before the survey, 12.9% of children of nontrained and 6.3% of children of trained WDA leaders, respectively, had diarrhea. During the same period, 15.8% and 6.3% of children of nontrained and trained WDA leaders had a fever and 21.1% and 13.5% had cough. Treatment seeking for a sick child was significantly different between trained WDA leaders (71%) and nontrained WDA leaders (43%).

5.4 Knowledge and practices of WDA network members

We assessed the knowledge and behavior of members of networks led by trained and nontrained WDA leaders using indicators related to WASH, DPC, and family health.

5.4.1 Water, sanitation, and hygiene

Trained WDA leaders' network members were more knowledgeable on several knowledge items than their counterparts led by nontrained leaders. No significant differences existed on most knowledge items related to CBT content for WDA leaders. These variabilities in the differences between network members with trained and nontrained leaders were observed in all areas of WASH-related knowledge, including the benefits of toilets, criteria for a healthy home, and methods for preventing insect and rodent infestation (Table 12).

Table 12. WASH-related knowledge among WDA network members with trained and nontrained leaders, February 2021

Knowledge items	Intervention WDA network members, N = 911, n (%)	Comparison WDA network members, N = 1825, n (%)	Total (N = 2736)	P-value
Benefits of using toilets				
Prevent communicable diseases	707 (77.6)	1,326 (72.7)	2,033 (74.3)	0.005
Prevent water and food poisoning	324 (35.6)	609 (33.4)	933 (34.1)	0.25
Protect environment from bad odors and pathogens	609 (66.8)	1,165 (63.8)	1,774 (64.8)	0.12
Criteria for healthy home				
Free from pathogens	597 (65.5)	1,105 (60.5)	1,702 (62.2)	0.011
Does not cause physical harm	225 (24.7)	439 (24.1)	664 (24.3)	0.71
Reduces social and psychological stress	169 (18.6)	343 (18.8)	512 (18.7)	0.88
Has shelf for foods and water utensils	363 (39.8)	709 (38.8)	1,072 (39.2)	0.61
Home should be far from toilets/separate/kitchen or pets	575 (63.1)	1,029 (56.4)	1,604 (58.6)	<0.001
Methods of preventing insect and rodent infestations				
Keep environment clean and dry and use sewage disposal	536 (58.8)	946 (51.8)	1,482 (54.2)	<0.001
Eliminate insect breeding areas	297 (32.6)	542 (29.7)	839 (30.7)	0.12
Eliminate predatory animals	549 (60.3)	1,167 (63.9)	1,716 (62.7)	0.061
Keep food away from such places	241 (26.5)	493 (27.0)	734 (26.8)	0.76
Adequate quality & safe water supply	95 (10.4)	155 (8.5)	250 (9.1)	0.098

Note: *P-values: Pearson chi-squared or Fisher exact test

Network members with trained WDA leaders better performed some of the recommended WASH-related practices than those with nontrained WDA leaders. For instance, open defecation was less common among those with trained WDA leaders (20.3 %) than among those with nontrained leaders

(30.5%). Network members with trained WDA leaders were more likely to participate in WASH campaigns (17.5% vs. 10.1%), to treat water (12.9% vs. 9.5%), and to practice handwashing before food preparation and after using toilets ($p \leq 0.05$). However, there was no statistically significant difference in other WASH-related practices, including handwashing practices during other critical times, ownership of improved sanitation facilities, availability of water for handwashing, separation of kitchen from living rooms, and the use of solid and liquid waste disposal systems (Table 13).

Table 13. WASH-related practices of WDA network members, February 2021

WASH-related practice items	Intervention WDA network members, N = 911, n (%)	Comparison WDA network members N = 1825, n (%)	Total (N = 2736)	P-value
Handwashing practices (critical times)				
Before preparing foods	762 (83.6)	1,477 (80.9)	2,239 (81.8)	0.022
While preparing foods	613 (67.3)	1,225 (67.1)	1,838 (67.2)	0.96
After preparing foods	666 (73.10)	1,274 (69.8)	1,940 (70.9)	0.073
After using toilet	645 (70.8%)	1,207 (66.1)	1,852 (67.7)	0.014
After cleaning a child who used toilet	241 (26.5)	458 (25.1)	699 (25.5)	0.44
After blowing nose, coughing, or sneezing	135 (14.8)	272 (14.9)	407 (14.9)	0.67
After touching animal, animal feed, or animal waste	312 (34.2)	670 (36.7)	982 (35.9)	0.18
After handling animal food or treatments	267 (29.3%)	537 (29.4%)	804 (29.4)	0.45
Toilet ownership and use				
Open defecation	185 (20.3)	557 (30.5)	742 (27.1)	<0.001
Pit latrine without slab	640 (70.3)	1,153 (63.2)	1,793 (65.5)	
Pit latrine with slab	77 (8.5)	108 (5.9)	185 (6.8)	
Other	9 (1.0)	7 (0.4)	16 (0.6)	
Improved sanitation facilities	80 (11.0)	110 (8.7)	190 (9.5)	0.086
Access to water and water treatment				
Water available for handwashing in compound	211 (27.4)	485 (24.7)		0.13
Treat water at household level	116 (12.9)	169 (9.5)	285 (10.7)	0.007
Adequate water treatment	93 (80.2)	114 (67.5)	207 (72.6)	0.018
Add chlorine or bleach	65 (56.0)	77 (45.6)	142 (49.8)	0.008
Time to fetch water, in minutes, m (SD)	30.1 (21.7)	34.0 (35.0)	32.7 (31.3)	0.006
Improved source of drinking water	805 (88.4)	1,451 (79.5)	2,256 (82.5)	<0.001
Household WASH facilities				
Water is available for handwashing in the house	232 (25.5)	464 (25.4)	696 (25.4)	0.98
Has shelf for storing feeding utensils	653 (71.7)	1,173 (64.3)	1,826 (66.7)	0.002
Feeding utensils covered and protected from flies	656 (72.0)	1,174 (64.3)	1,830 (66.9)	0.001

WASH-related practice items	Intervention WDA network members, N = 911, n (%)	Comparison WDA network members N = 1825, n (%)	Total (N = 2736)	P-value
Separate kitchen	700 (76.8)	1,350 (74.0)	2,050 (74.9)	0.10
Separate sleeping room	615 (67.5)	1,154 (63.2)	1,769 (64.7)	0.027
Separate room for livestock	685 (75.2)	1,364 (74.7)	2,049 (74.9)	0.80
Participated in WASH campaign	159 (17.5)	184 (10.1)	343 (12.5)	<0.001
Functional solid waste disposal	255 (28.0)	449 (24.6)	704 (25.7)	0.056
Functional liquid disposal	129 (14.2)	248 (13.6)	377 (13.8)	0.68

Note: *P-values: Pearson chi-squared or Fisher exact test

5.4.2 Disease prevention and control

We found that WDA network members with trained leaders were more likely to have heard about communicable diseases, including TB and HIV/AIDS, than their counterparts with nontrained leaders. Contrary to our expectations, WDA network members with trained leaders were more likely than those with nontrained leaders to mistakenly believe HIV can be acquired from mosquitos. There was also no significant difference between WDA network members with trained and nontrained leaders in relation to more detailed prevention- and care-related questions, including the curability of TB, availability of ARV treatment for HIV, and major misconceptions about HIV (Table 14).

Table 14. Awareness of, knowledge of, and practices related to TB and HIV/AIDS among WDA network members with trained and nontrained leaders, February 2021

	Intervention network members, N = 911 n (%)	Comparison network members N = 1,825 n (%)	Total (N = 2376)	P-value
Knowledge and perception of TB				
Heard about TB	769 (84.4)	1,417 (77.6)	2,186 (79.9)	<0.001
Know TB can be cured	640 (83.2)	1,169 (82.5)	1,809 (82.8)	0.67
Knowledge and perception of HIV/AIDS				
Heard of HIV/AIDS	879 (96.5)	1,640 (89.9)	2,519 (92.1)	<0.001
Believe people can reduce their chance of getting HIV by having just one uninfected sex partner who has no other sex partners	714 (81.2)	1,213 (74.0)	1,927 (76.5)	<0.001
Believe people can get HIV from mosquito bites	304 (34.6)	503 (30.7)	807 (32.0)	0.045
Believe people can reduce chance of getting HIV by using a condom every time they have sex	631 (71.8)	1,045 (63.7)	1,676 (66.5)	<0.001
Believe people can get HIV by sharing food with a person who has HIV	736 (83.7)	1,381 (84.2)	2,117 (84.0)	0.76
Believe it possible for a healthy-looking person to have HIV	716 (81.5)	1,221 (74.5)	1,937 (76.9)	<0.001
Heard of ARVs to treat HIV	603 (68.6)	1,092 (66.6)	1,695 (67.3)	0.30

	Intervention network members, N = 911 n (%)	Comparison network members N = 1,825 n (%)	Total (N = 2376)	P-value
Know about special medicines a doctor or a nurse can give a woman infected with HIV to reduce the risk of transmission to the baby	617 (70.2)	993 (60.5)	1,610 (63.9)	<0.001

Note: *P-values: Pearson chi-squared or Fisher exact test; ARV: Anti-retroviral therapy

Similarly, awareness about NCDs, including diabetes, hypertension, and cervical cancer, was higher among network members with trained WDA leaders than among those with nontrained leaders. However, knowledge of methods for prevention of cervical cancer and treatment-seeking behavior for sick household members in general were not statistically different between the two groups (Table 15).

Table 15. Awareness of, knowledge of, and practices related to NCDs, adulthood illnesses, and treatment-seeking behaviors among WDA network members with trained and nontrained WDA leaders, February 2021

	Intervention WDA network members, n (%)	Comparison WDA network members, n (%)	Total n (%)	P-value
Knowledge and perception of NCDs				
Heard of NCDs (hypertension, DM, cancer)	665 (73.0)	1,238 (67.8)	1,903 (69.6)	0.006
Primary sources of information about NCDs				
Health extension worker	275 (41.4)	284 (22.9)	559 (29.4)	<0.001
WDA leaders	62 (9.3)	84 (6.8)	146 (7.7)	
Health center or hospital staff	89 (13.4)	249 (20.1)	338 (17.8)	
Media	137 (20.6)	345 (27.9)	482 (25.3)	
Other	102 (15.3)	276 (22.3)	378 (19.9)	
Heard about cervical cancer	436 (47.9)	723 (39.6)	1,159 (42.4)	<0.001
Knowledge of cervical cancer prevention methods				
Avoid unprotected sexual intercourse	96 (22.0)	150 (20.7)	246 (21.2)	0.61
Vaccination	182 (41.7)	280 (38.7)	462 (39.9)	0.31
Screening	244 (56.0)	440 (60.9)	684 (59.0)	0.10
Treatment-seeking behaviors (household members)				
At least one household member was sick in the last 12 months	281 (30.8)	545 (29.9)	826 (30.2)	0.60
Treatment sought for a sick household member (most recent)	239 (85.1)	434 (79.6)	673 (81.5)	0.057
Place where treatment was sought for a sick household member				
Health center/health post	140 (58.8)	236 (54.4)	376 (56.0)	0.70
Hospital	57 (23.9)	110 (25.3)	167 (24.9)	
Private clinic	40 (16.8)	86 (19.8)	126 (18.8)	
Other	1 (0.4)	2 (0.5)	3 (0.4)	

Note: *P-values: Pearson chi-squared or Fisher exact test NCD: Non-Communicable Diseases; DM; Diabetes Mellitus

5.4.3 Family health

Family planning

Knowledge about commonly used FP methods, including injectables, implants, and OCPs, was high among both groups of WDA network members. For the remaining methods, WDA network members with trained leaders were slightly more knowledgeable than those with nontrained leaders (Figure 15).

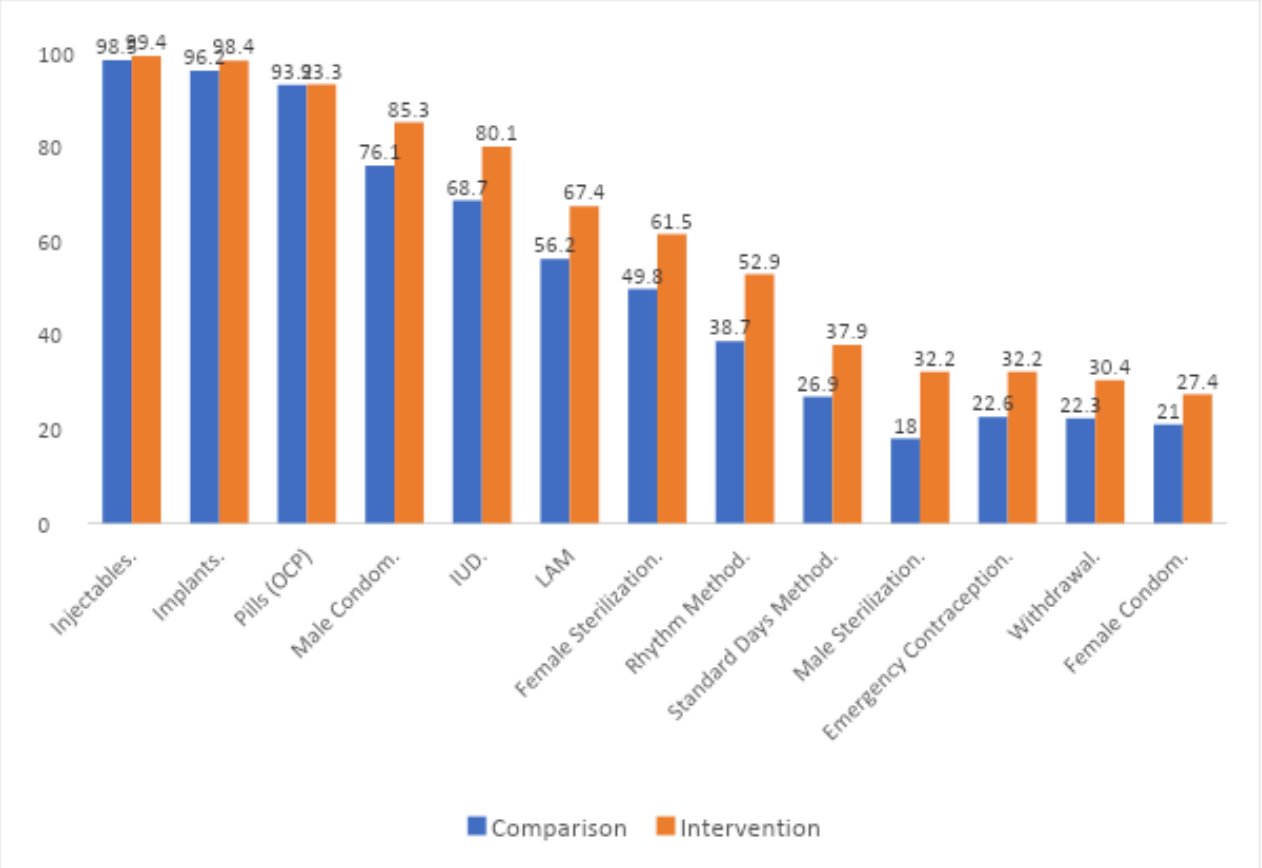


Figure 15. Knowledge of FP methods among WDA network members by group

Having ever used and current use of FP methods were not statistically different between network members with trained WDA leaders and nontrained WDA leaders. The prevalence of mCPR use was 39.7% among network members with trained WDA leaders and 39.4 % among network members with nontrained WDA leaders. The use of LARC was 26.5% among network members with trained WDA leaders and 29.9% among network members with nontrained WDA leaders (Table 16).

Table 16. Reproductive history and use of FP methods among WDA network members with trained and nontrained leaders, February 2021

Characteristics	Intervention WDA network members, n (%)	Comparison WDA network members n (%)	Total n (%)	P-value
Reproductive history				
Ever been pregnant	605 (94.7)	1,051 (94.3)	1,656 (94.4)	0.71
Currently pregnant	57 (9.4)	98 (9.3)	155 (9.4)	0.45
Current pregnancy not wanted	12 (21.0)	16 (16.0)	28 (18.1)	0.46
Most recent pregnancy				
Wanted	511 (84.5)	896 (85.3)	1,407 (85.0)	0.36
No, wanted later	56 (9.3)	116 (11.0)	172 (10.4)	
Not wanted at all	38 (6.3)	39 (3.7)	77 (4.6)	
Main sources of information about FP				
HEWs	408 (65.7)	479 (44.9)	887 (52.6)	<0.001
Health professionals	73 (11.8)	206 (19.3)	279 (16.5)	
Other sources	140 (22.5)	381 (35.7)	521 (30.9)	
Practices: FP				
Ever used FP (any FP methods)	264 (45.4)	445 (43.8)	709 (44.3)	0.53
Used FP in the last 12 months (any FP methods)	288 (45.1)	460 (41.3)	748 (42.6)	0.12
Current use, modern CPR	254 (39.7)	439 (39.4)	693 (39.5)	0.88
Current use, LARC	70 (26.5)	133 (29.9)	203 (28.6)	0.34
Source of FP				
Health center	85 (32.2)	221 (49.7)	306 (43.2)	<0.001
Health post	131 (47.6)	170 (38.2)	301 (42.5)	
Other	48 (18.2)	54 (12.1)	102 (14.4)	

Note: *P-values: Pearson chi-squared or Fisher exact test

Injectable contraceptives, followed by implants, constituted the most used FP methods among WDA network members in both the intervention and comparison groups. The share of LARC among current users of contraceptives is 30% in the intervention and 27% in the comparison group. The use of LARC (implants) was more common in Amhara and SNNPR than in Oromia (Figure 16).

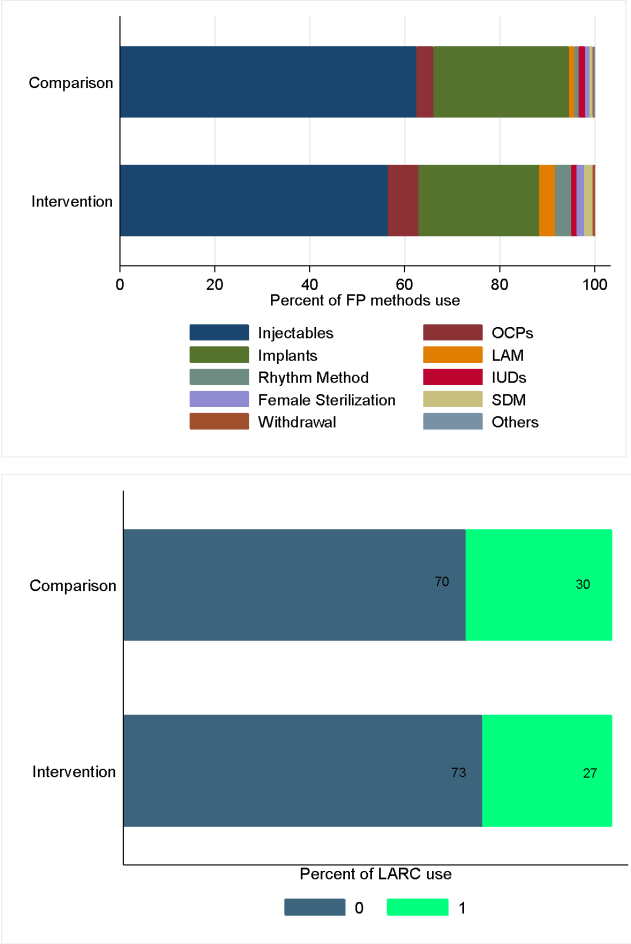


Figure 16. Contraceptive method mix among current FP user WDA network members with trained and nontrained leaders

Maternal health

Knowledge about the danger signs of pregnancy was better in the intervention than in the comparison group. Higher proportions of WDA leaders knew danger signs in pregnancy, including vaginal bleeding (75% vs. 64%), edema (26% vs. 22%), severe headache (34% vs. 27%) and convulsions (18% vs 10%) (Figure 17).

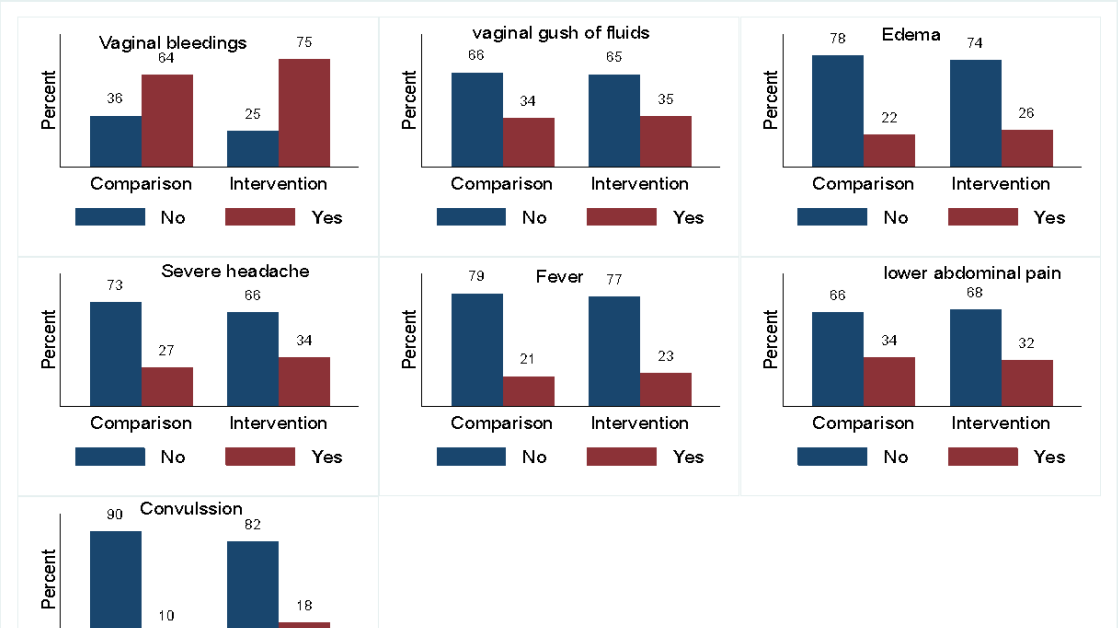


Figure 17. Awareness about danger signs of pregnancy among WDA network members with trained and nontrained leaders, February 2021

We also asked respondents about the recommended time for the first PNC visit. Responses showed that 32% of network members with trained WDA leaders and 27% of those with nontrained leaders believed the appropriate time for the first PNC visit to be within 24 hours after delivery.

During their latest pregnancy, 87.9% of WDA network members with trained leaders and 85.5% of network members with nontrained leaders received ANC. Coverage of at least four ANC visits during the most recent pregnancy was 56.4% in the intervention group and 42.9 % in the comparison group. Health facility delivery and postnatal checkup were also more widespread among network members with trained WDA leaders than those with nontrained leaders (Table 17).

Table 17. Use of maternal healthcare services among WDA network members with trained and nontrained leaders, February 2021

Characteristics	Intervention WDA network members, n (%)	Comparison WDA network members, n (%)	Total n (%)	P-value
Received ANC during last pregnancy	106 (89.8)	201 (86.6)	307 (87.7)	0.39
4+ ANC visits during last pregnancy	61 (57.5)	83 (41.3)	144 (46.9)	0.007
Place of ANC services				
Health center	71 (67.0)	136 (67.7)	207 (67.4)	0.58
Health post	23 (21.7)	49 (24.4)	72 (23.5)	
Hospital	12 (11.3)	16 (8.0)	28 (9.1)	
Place of birth for the last birth				
Home	279 (46.7)	576 (55.5)	855 (52.3)	<0.001
Health facility	319 (53.3)	461 (44.5)	780 (47.7)	
PNC visit for most recent birth in the last 12 months	69 (58.5)	129 (55.6)	198 (56.6)	0.61
Place of PNC services				
Healthcare facilities	58 (84.1)	116 (89.9)	174 (87.9)	0.23

Note: *P-values: Pearson chi-squared or Fisher exact test

Child vaccination

Child vaccination status was assessed among WDA network members' children aged 12–23 months using information from vaccination cards. Ownership of a child vaccination card was reported for 57.9% of the children in the intervention and 52.5% in the comparison group. Of these, interviewers could verify only 76.6% and 79.6%, respectively. Coverage of full basic vaccines was 63.4% in the intervention group and 48.1% in the comparison group. Vaccination coverage was highest for BCG, first and second doses of polio, pentavalent, and rotavirus vaccines (Figure 18).

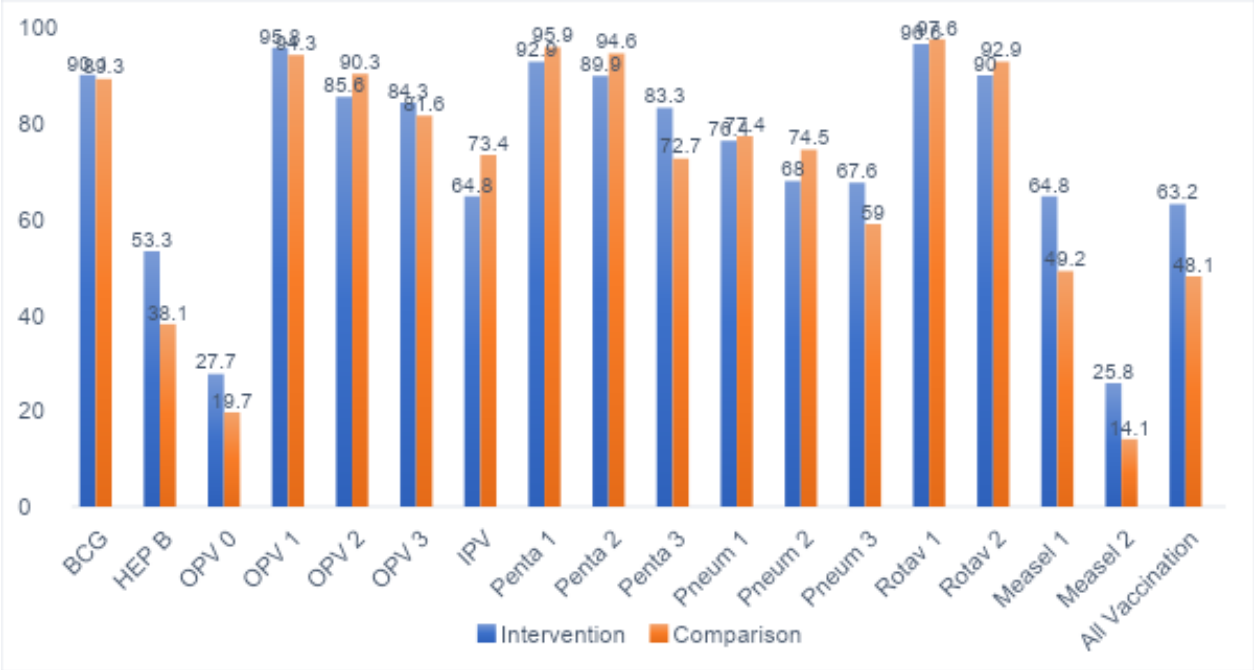


Figure 18. Vaccination coverage among children 12–23 months of WDA members with trained and nontrained leaders, February 2021

Childhood illnesses and treatment-seeking behavior

Eleven percent of children under five years of age in the intervention group and 12.7% in the comparison group had diarrhea in the two weeks before the survey. Additionally, 14.6% and 13.8% of children in the comparison and intervention groups had a fever and an almost equal proportion of children in the intervention group (16.3%) and in the comparison group (16%) had a cough in the same period. Treatment seeking for a sick child was 49.8%, ranging from 47% in the comparison group to 54.5% in the intervention group (Figure 19).

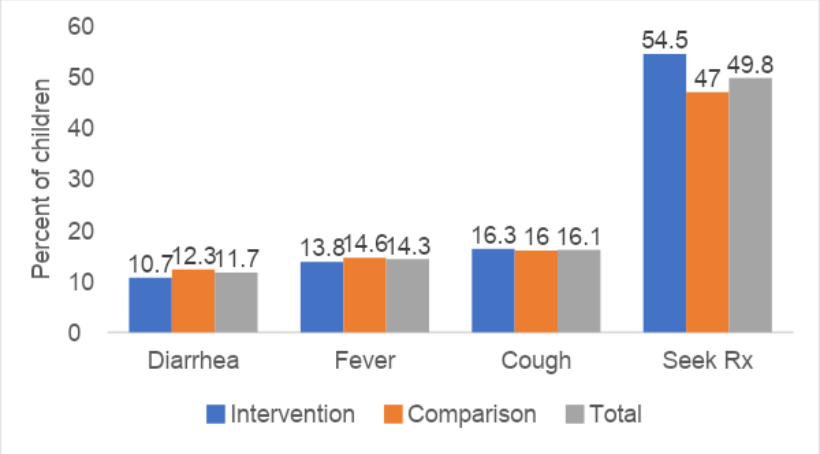


Figure 19. Childhood illnesses and treatment-seeking behaviors among WDA network members with trained and nontrained leaders, February 2021

5.5 Effect of CBT on HEP-related knowledge and behaviors among WDA leaders and their network members

In addition to the reaction of WDA leaders to CBT, measured in terms of the satisfaction of trainees with different aspects of the training and presented under Section 2, CBT was evaluated based on learning- and behavior-level outcomes assessed among WDA leaders. Moreover, at the level of results, CBT was evaluated based on a comparison of HEP-related knowledge and behavior between network members of trained and nontrained WDA leaders.

Bivariate analysis revealed that exposure of WDA leaders to CBT is associated with increased knowledge and improved behavior regarding some of the targeted WASH, maternal and child health, and disease prevention and control interventions. Table 18 presents findings from doubly robust estimation methods on the effect of CBT on knowledge and behavior of WDA leaders and their network members.

The program achieved some the intended outcomes of CBT for both WDA leaders and network members and some for either the leaders or the members. For the remaining outcomes, no evidence suggests that CBT has any effect on WDA leaders or network members. Among 30 indicators assessed, CBT was found positively and significantly associated with 13 (43.3%) for WDA leaders and 19 (63.3%) for WDA members (Figure 20, Table 18).

		CBT has positive effect on WDA members	
		Yes	**No**
CBT has positive effect on WDA leader	**Yes**	1. Knowledge about FP	1. Knowledge about WASH
		2. Health facility delivery	2. Functional solid waste disposal pit
		3. Household treatment of drinking water	3. Use of soap
		4. Covering food and protecting it from flies	4. Perception that TB can be cured
		5. Improved sources of drinking water	
		6. Participation in WASH campaigns	
		7. Knowledge about HIV/AIDS	
		8. Heard of NCDs	
		9. Heard of cervical cancer	
	No	1. FP use in the last 12 months	1. Use of LARC
		2. Modern contraceptive prevalence rate	2. PNC coverage
		3. ANC (4 plus) visits	3. Wash hands at critical times
		4. Coverage of all basic child vaccines	4. Household has a separate kitchen
		5. Avoid open defecation	5. Household has a separate sleeping room
		6. Practice adequate water treatment	6. Household has a separate room for animals
		7. Clean and cover feeding utensils	7. Improved sanitation facilities
		8. Wash fruits and vegetables	
		9. Functional liquid disposal system	
		10. Seek treatment for sickness	



Figure 20. Indicators of aspects of CBT that affected WDA leaders and network members

On average, WDA leaders who received CBT knew 2.4 more FP methods than those who did not receive the training (β = 2.408, 95% CI [1.97, 2.841]). Similarly, the number of FP methods known by WDA members with trained leaders was higher by 1.3 than that among WDA members with nontrained leaders (β = 1.255, 95% CI [1.03, 1.48]). On the other hand, CBT was not significantly associated with the use of modern contraceptives among WDA leaders (β = –0.023, 95% CI [–0.103, 0.056]), while it showed a weak yet statistically significant association with that of WDA members. Network members with trained WDA leaders had a 4.8% higher mCPR than those with nontrained WDA leaders (β = 0.0482, 95% CI [0.012, 0.084]). Table 18 presents adjusted coefficients of association between CBT for WDA leaders and a number of health indicators among WDA leaders and their network members, adjusted for the potential confounding effects of socioeconomic status, age of women, educational status, agricultural land size, head of the household, region, TLU, and membership in CBHI.

Table 18. Effect of CBT on knowledge and practices of WDA leaders and network members, February 2021.

Outcomes of interests	**Effect of CBT on WDA leaders**			**Effect of CBT on network members**		
	Adjusted Doubly Robust estimation, IPWRA (ATT)			Adjusted Doubly robust estimation, IPWRA (ATT)		
	Coef.	SE	[95% CI]	Coef.	SE	[95% CI]
Maternal and child health						
Number of FP methods known **	2.408	.220	1.97, 2.841	1.255	0.115	1.03, 1.48
FP use (12 months) **	.024	.042	-.058, .106	.0718	.019	.035, .109
mCRP**	-.023	.040	-.103, .056	.0482	.018	.012, .084
LARC	.032	.076	-.118, .1811	-.0365	.036	-.103, .030
ANC (4 plus) **	.185	.128	-.065, .437	.166	.058	.05, .281
Deliver at health facilities**	.100	.044	.013, .187	.101	.020	.061 .141
Had PNC for latest child	.218	.121	-.018, .455	.042	.055	-.065, .149
Child received all basic vaccination**	-.104	.141	-.382 .172	.151	.061	.031, .271
WASH						
HH treat water**	.108	.035	.041, .175	.034	.013	.008, .060
HH practice adequate water Rx**	-.048	.078	-.201,.104	.127	.052	.026, .228
Wash hands (critical time index)	.081	.093	-.102,.264	.022	.040	-.057, .101
Knowledge about WASH*	.184	.091	.005,.364	-.002	.039	-.078, .074
Practice open defecation**	-.063	.034	-.132, .006	-.102	.017	-.136, -.068
HH have a separate kitchen	-.028	.036	-.099, .043	.020	.017	-.013, .053
HH cover food from flies	.034	.038	.041, 0.11	.078	.018	.04,.113
HH clean and cover feeding utensils	-.003	.037	-.078,.071	.102	.018	.066,.138
HH Wash fruits and vegetables	.024	.025	-.025, .074	.020	.013	.034, .092
HH have a separate sleeping room	-.0269	.041	-.108,.054	.043	.019	-.005, .08
HH has separate room for animals	-.005	.039	-.083, .073	-.007	.018	-.043, .029
HH has functional SW disposal*	.142	.044	.055, .229	.007	.017	-.027, .042
HH has functional liquid disposal*	.026	.058	-.089, .142	.076	.034	.01, .142

Improved sources of drinking water**	.086	0.029	.029,0.144	.089	.014	.061,.116
Improved sanitation	.022	0.033	-.044,.088	.023	.014	-.004, .051
Use of soap	.085	0.041	.006, .166	.027	.016	-.004,.058
HH Participate in WASH campaign**	0.163	0.049	.083, .243	.063	.015	.034, .092
Disease Prevention and Control						
Knowledge score about HIV/AIDS**	.350	.107	.14, .56	.306	.060	.188, .424
TB can be cured (perception)	.068	.031	.009, .128	.007	.016	-.026, .04
Seek treatment for illness **	.068	.050	-.032, .167	.054	.027	0.001, .108
Ever heard of NCDs**	.103	.033	.037, .169	.052	.018	.016, .087
Ever heard of cervical Cancer**	.243	.039	.165, .321	.082	.020	.043, .122

Note: Models are adjusted for socio economic status, age of women, educational status, agriculture land size, head of the household, region, TLU, and CBHI membership.; NCDs: Non-communicable Diseases

Note: Models are adjusted for socioeconomic status, age of women, educational status, agricultural land size, head of the household, region, TLU, and CBHI membership.

Because of variabilities in the implementation of CBT across settings, we further investigated if reported length of training hours, competency assessment, and certification could be associated with improved knowledge and practices of WDA leaders. We regressed length of training hours, competency assessment, and certification status along with potential confounders on maternal and child health outcomes, including mCPR, LARC use, and health facility delivery, and WASH related outcomes, including improved sanitation, water treatment, practice of open defecation, and handwashing during critical times. Our findings from multiple regression analysis show that, overall, there is no statistically significant association between these characteristics of CBT and the health outcomes (Table 19).

Table 19. Association of training hours, competency assessment, and certification with health outcomes

		Covariates		
		Length of training hours	Received competency assessment	Certified
Outcomes	mCPR	–.024 (–.058, .010)	.646 (–.516, 1.809)	.299 (–.993, 1.591)
	LARC	–.060 (–.174, .054)	1.707 (–4.402, 7.816)	–4.606 (–11.90, 2.695)
	Healthcare facilities delivery	–.028 (–.069, .012)	.663 (–.704, 2.030)	1.145 (–.3776, 2.668)
	Improved source of drinking water	.111 (.027, .196) **	–.417 (–2.877, 2.042)	–1.612 (–4.256, 1.032)
	Improved sanitation	.014 (–.048, .076)	.644 (–1.352, 2.640)	.757 (–1.323, 2.838)
	Treat water	.009 (–.033, .053)	–.315 (–1.785, 1.154)	.051 (–1.422, 1.525)
	HH practices open defection	.047 (–.035, .130)	.685 (–1.511, 2.882)	NA
	Handwashing during critical time	.016 (.001, .031) **	–.130 (–.659, .397)	.265 (–.326, .856)

Note: Adjusted for socioeconomic status, age of women, educational status, agriculture size, head of the household, region, TLU, Participation in member community-based health insurance (CBHI).; mCPR: modern Contraceptive Prevalence Rate

5.6 The effect of CBT on the functionality of WDA structures

In the 12 months before the survey, trained WDA leaders conducted a higher number of meetings (M = 7.34) than nontrained WDA leaders (M = 4.92; p = 0.027). The proportion of WDA leaders who conducted at least one meeting in the preceding three months was higher for the trained group (54.1%) than for the nontrained group (35.4%; P = 0.01; Table 20).

Table 20. Functionality of WDA structures led by trained and nontrained WDA leaders, February 2021

Characteristics	Intervention (trained WDA leaders, N = 187), n (%)	Comparison (nontrained WDA leaders, N = 374) n (%)	Total (N = 561) n (%)	P-value
Role as WDA leaders				
1-5 network WDA leaders	158 (84.5)	312 (83.4)	470 (83.8)	0.75
1-30 network WDA leaders	29 (15.5)	62 (16.6)	91 (16.2)	
Months of service as WDA leader, mean (SD)	37.08 (25.63)	41.05 (43.57)	39.72 (38.55)	0.25
Average number of network members, mean (SD)	5.31 (3.32)	5.45 (3.51)	5.40 (3.44)	0.67
Know members (yes)	174 (93.0)	322 (86.1)	496 (88.4)	0.015
WDA leader mentioned all members of her network	159 (91.4)	282 (87.6)	441 (88.9)	0.20
WDA meetings				
At least one meeting during the last 12 months	111 (59.4)	79 (21.1)	190 (33.87)	<0.001
Number of meetings conducted in the last 12 months	7.34 (7.84)	4.92 (6.70)	6.33 (7.47)	0.027
At least one meeting during the last three months	60 (32.1)	28 (7.5)	88 (15.7)	<0.001
Meeting with specific groups during the last 12 months				
Met with 1-5 network members	92 (49.2)	72 (19.3)	164 (29.2)	<0.001
Met with 1-30 WDA leaders	39 (20.9)	21 (5.6)	60 (10.7)	<0.001
Met with 1-30 network members	24 (12.8)	7 (1.9)	31 (5.5)	<0.001
Met with HEWs	58 (31.0)	28 (7.5)	86 (15.3)	<0.001
Assisted HEWs	136 (72.7)	118 (31.6)	254 (45.3)	<0.001
Participated in plan development in the last 12 months	92 (49.2)	72 (19.3)	164 (29.2)	<0.001
Created performance report in the last 12 months	39 (20.9)	21 (5.6)	60 (10.7)	<0.001
Disseminated health message in the last 12 months	89 (23.8)	126 (67.4)	215 (38.3)	<0.001
Participated in non-health activities	58 (31.0)	28 (7.5)	86 (15.3)	<0.001
Conducted home visits in the last 12 months	109 (58.3)	83 (22.2)	192 (34.2)	<0.001
Conducted outreach activities	69 (36.9)	50 (13.4)	119 (21.2)	<0.001
Participated in campaign	78 (41.7)	56 (15.0)	134 (23.9)	<0.001
Community mobilization	67 (35.8)	54 (14.4)	121 (21.6)	<0.001
Used family health card	116 (62.0)	50 (13.4)	166 (29.6)	<0.001
Reported any challenges	113 (60.4)	80 (21.4)	193 (34.4)	<0.001
WDA implementation intensity score, mean (SD)	4.72 (2.74)	1.58(2.26)		<0.001

Note: *P-value: Pearson chi-squared or Fisher exact test

The overall mean score of WDA functionality was 2.63 (SD = 2,84), and we found the functionality of WDA structures led by CBT-trained leaders higher (M = 4.72, SD = 2.74) than those led by nontrained ones (M = 1.58, SD = 2.26; P = 0.001). The mean score of WDA functionality varies by region. It is highest in Amhara (M = 3.38, SD = 2.74), followed by SNNPR (M = 2.5, SD = 2.94), and lowest in Oromia (M = 1.94, SD = 2.70; Figure 21).

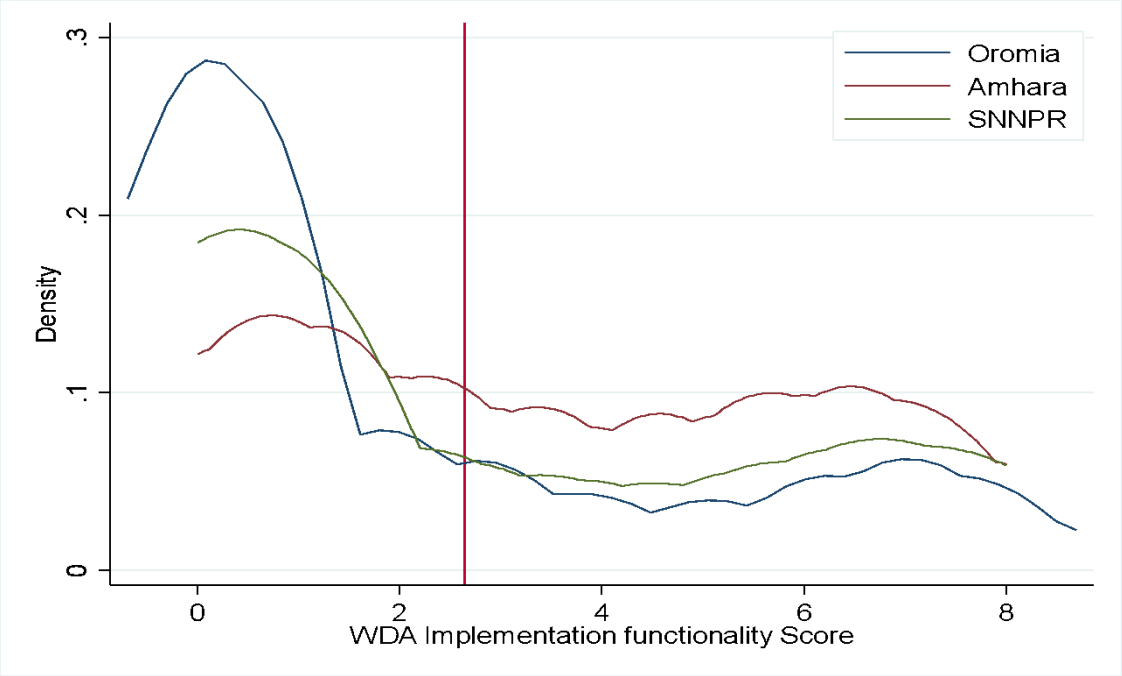


Figure 21. WDA implementation functionality score by region, February 2021

We used doubly robust estimation to analyze the association between provision of CBT for WDA leaders and the functionality of WDA structures. Our analysis showed that, after adjusting for potential confounders, including socioeconomic status, age of women, educational status, agricultural land size, head of the household, region, TLU, and membership in CBHI, WDA networks led by trained leaders have a 2.555 higher WDA functionality index than those led by nontrained WDA leaders (β = –2.255, 95% CI [1.898, 2.612]).

Most trainees who participated in key informant interviews also reported that they were involved in collaborative work with HEWs. Moreover, after participating in CBT, they began sharing the knowledge and skills they gained from the training with their network members and with other community members through formal meetings and during community gatherings and other social activities. A WDA leader from SNNPR said,

> I share the knowledge I have with the 1-5 network group and with nonmembers when we meet for coffee ceremony and women's edir [burial association groups]. These women also share their knowledge with other women. Previously, women were shy. However, after attending CBT women have become self-confident, and they now can discuss their problems in their home environment. (WDA leader, Meskan woreda, Gurage Zone, SNNPR)

Another WDA leader also said the following:

> I stopped my work and attended this training. After the training, I practically applied it to my daily life. I also trained my team members. I trained my team members on environmental hygiene and other healthcare packages. I told them to use FP methods so that their children will grow healthier. (WDA leader, Shirka woreda, Arsi zone, Oromia)

The HEP coordinators also appreciated the role of CBT in increasing the extent of engagement of WDA leaders in awareness creation:

> During our visit, we have seen that WDA leaders organized a meeting and shared the information with their network members. Before the training, many of the women were used to giving birth at home. However, home delivery becomes shameful after the CBT. (HEP Coordinator, Shirka wereda, Arsi zone, Oromia)

In addition to maternal health, trained WDA leaders started awareness creation on communicable disease and contributed to disease prevention. One HEW said the following during an in-depth interview:

> If you see communicable diseases, some of community members could not identify the difference between tuberculosis and a cough before the training. Due to this, patients were living with healthy family members under the same roof without taking care and protection. Nevertheless, trained WDA leaders advised these patients to go to health centers by creating awareness about the disease. (HEW, Sinana woreda, Bale zone, Oromia)

HEW and HEP coordinators reported that trained WDA leaders assisted HEWs in information dissemination, identification of individuals eligible for different services, and supporting HEWs in different ways. Informants uniformly linked CBT with improvements in these functions of WDA networks.

> If there is vaccination program or a meeting, HEWs will call trained WDA leaders and inform them to gather their network members and their neighbors. If a mother goes into labor, they will call an ambulance. After CBT, their support to HEWs in these activities and the like is significant. (HEP-coordinator, Meskan woreda, Gurage zone, SNNPR)

> It [CBT] reduces the load or activities, implementing the planned activities as required, and to share health information easily. The neighbors have understood and learned from each other and there will be a wide range of change in this village [because of CBT]. (HEW, Shirka woreda, Arsi zone, Oromia)

In addition to its contributions in increasing the engagement of WDA leaders in the HEPs' activities, CBT was also reported to have improved the understanding of HEWs about the capacities of WDA leaders in their communities, facilitating improved relationships and engagement with them. A HEW from Bale zone in Oromia described the benefits of CBT:

> I am satisfied with the CBT because of three things. First, it enhanced my relationship with the community.... Third, I had the chance to know all women's levels of understanding and knowledge and their gaps, which enabled me to identify the strong women who can change the community as role models. It also reduced my workload. When I want to pass on health messages for the community, what I need to do is just to call trained WDA leaders and

communicate with them. They can easily inform their members. (HEW, Sinana woreda, Bale zone, Oromia)

5.7 Facilitators of and barriers to CBT

5.7.1 Performance of CBT in study woredas

Performance reports from WoHOs show that competency assessment and certification of WDA leaders were suboptimal. According to reports for the past three years, 62.7% of the WDA structures were available and functional. Only 13.5% of WDA leaders have received CBT, but this varies by region. We have found a vast gap between provision of the training and administration of the assessment. Only 7.2% of WDA leaders had been assessed, and 6.9% were certified (Figure 22).

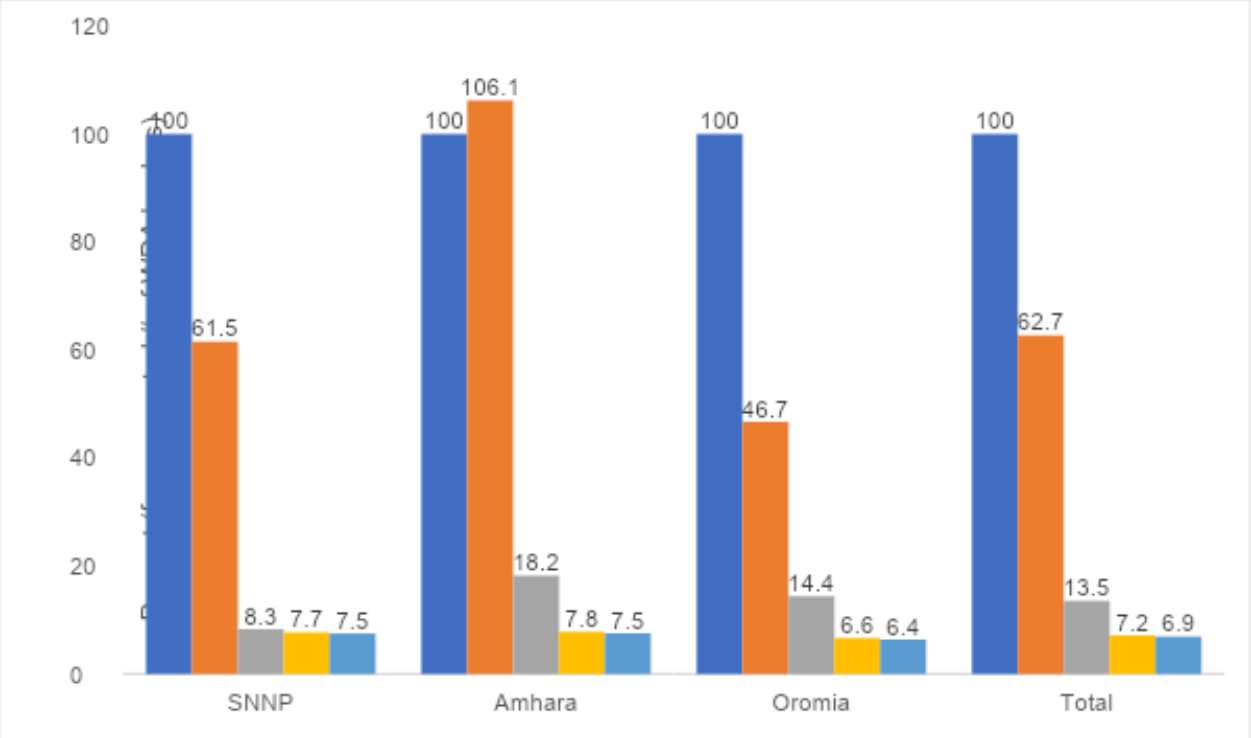


Figure 22. Implementation status of CBT activities among sample kebeles by region

5.7.2 Facilitators of and barriers to the performance of CBT

Using thematic analysis of the qualitative data, we identified facilitators and barriers that influenced the performance of CBT. The analysis identified three groups of factors:

1. Micro-level factors, which are related to WDA leaders' or trainers' characteristics and conditions, including literacy status and economic hardship
2. Meso-level factors, which operate in community and environmental contexts, such as sociocultural factors, including sociocultural norms, values, practices and beliefs, gender roles or household responsibilities, and travel distances
3. Macro-level factors operating in health-system policy and practices, including recruitment or selection of WDA leaders, kebele processes, manual preparation, incentives, coordination among stakeholders, and political commitment (see Figure 23)

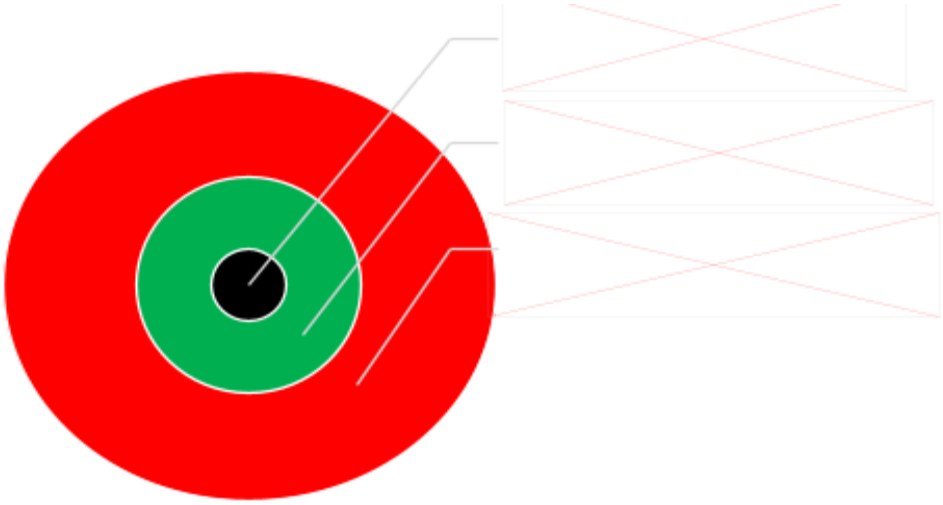


Figure 23. Categories of barriers to and facilitators of CBT implementation, February 2021

Micro-level factors: Factors related to WDA leaders and trainers

WDA leaders' characteristics, such as illiteracy, lack of prior knowledge, beliefs and attitudes, and living conditions (*e.g.*, economic hardship and living in a single room), and variabilities in these characteristics hindered the implementation and effectiveness of CBT. On the other hand, trainer-related factors such as motivation, satisfaction, readiness to deliver the course, and competency positively influenced implementation. One of the trainers explained this situation by saying,

> We have taken more time to discuss FP. The trainees had some reservations about issues related to FP. There was confusion and disagreements while we were teaching them about the available methods. They probed a lot on *how* and *why* questions of FP methods. Because of this, it took more time to explain the details of family planning for them. (HEW, Bati woreda, Oromia special zone, Amhara)

> All trainees do not have the same level of understanding; some were fast, and others took a longer time to understand. To bring them on the same page, HEWs had to explain some topics repeatedly, which took them more time to cover. (HEP-coordinator, Meskan woreda, Gurage zone, SNNPR)

WDA leaders' positive attitude and altruism could facilitate the success of the program. Positive relationships between trainers and WDA leaders have been among the motivators for WDA leaders to complete their training sessions. WDA leaders described this as follows:

> It changed many things. I give support as much as possible for those staying at their home. We motivated and supported other women using resource we have at hand. We supported and advised those faced with problems at their homes. Those not using toilets, we encourage to have one and use the latrine, and clean their kitchen and environment. (WDA leader, Mojana-Wodera Woreda, South Wollo zone, Amhara)

> I am telling you frankly the approach of the trainer was the satisfying part because our relationship with her was like a friend, which made the training more attractive. We were

following the training with interest. We engaged proactively and eagerly until the next training session. (WDA leader, Tiyo woreda, Arsi zone, Oromia)

Preparatory activities in terms of the training of trainers seem to have substantially contributed to building the competency of trainers themselves, which positively influences the implementation of CBT. HEWs' competency in providing the training was one of the facilitating factors:

> Two trainers who gave the training are training of trainers (TOT) certified. Both trainers had good skills and capacity because they were first trained on the modules themselves and they had previous experience on the program. Their approach to proving the training, clarity of ideas, and assessment method was very, very good. (HEP-coordinator, Tiyo woreda, Arsi zone, Oromia)

Meso-level factors: Community and environmental contexts

Several factors operating at the community and environment levels affected the implementation and effectiveness of CBT. The WDA leaders having many responsibilities with their household chores compromised the uptake of the training. Whenever WDA leaders attended the training, it became a source of dissatisfaction for them, as they had to compromise.

Inconvenient training venues were among the environmental factors negatively affecting the program. Some training sessions were facilitated under trees or in open spaces where trainers and trainees were exposed to sun and heat. The location of training centers was also cited as one of the main challenges to participate in the training. Some reported that the training center was far from trainees' homes.

> There were mothers who came for the training after traveling for more than an hour. The training was conducted in summer, and the road was so muddy. When they reached the training place, they got tired and felt sleepy. Some mothers needed coffee, but there was no refreshment. (HEW, Enamor Ener woreda, Gurage zone, SNNPR)

Whereas some said the venues were attractive, others claimed they were inadequate or too far away. Both trainers and trainees reported the difficulty of securing appropriate venues that can accommodate all trainees and that are comfortable and closer to trainees' homes. Some participants reported that they used schools, government offices, and farmers' training centers.

> The training was given in one of the schools. The training was very attractive, and we did not face any difficulty. (HEW, Mojana-Wodera Woreda, South Wollo zone, Amhara)

> The venue was not bad, but the room was too small to accommodate all trainees… It was not adequate. (HEW, Enamor Ener woreda, Gurage zone, SNNPR)

> We held the training in the health post compound, and we used a mat for sitting. Since the trainees were 25, the mat was not enough, indeed. The space was also small. (HEP coordinator, Meskan woreda, Gurage Zone, SNNPR)

The training seasons and schedules were sometimes inconvenient. While consulting the WDA leaders about the training schedule facilitates the program, providing the program while WDA leaders are busy with other household and farm activities may hinder implementation.

The preexistence of functional WDA structures and communities' acceptance of WDA leaders also emerged as important determinants of the implementation and effectiveness of CBT. WDA functionality affects the CBT program: where the WDA structure was functional, most WDA leaders performed better in conducting meetings with their network members, assisting HEWs, providing home-to-home education services, and promoting healthcare service use and environmental sanitation activities. The community's acceptance facilitates the success of the program. It also improves acceptance of HEWs.

5.7.3 Macro-level factors operating at health system policy and practices

At this level, resource allocation, selection of WDA leaders, stakeholder involvement and coordination, recognition, and the training design (methods) were important factors that operated as facilitators if they were adequate and appropriate and as barriers if they were not.

Resource allocation

Lack of refreshments and incentives during training sessions were commonly reported challenges for CBT implementation. According to trainers and training coordinators, these factors are linked to the inadequate allocation of resources for the facilitation of training. In addition, inadequate resource allocation for the program has led, in some places, to inadequate supervision and insufficient training manuals, both of which were common barriers to the implementation of CBT.

A HEW linked this with inadequate budget allocation:

> Frankly speaking, the training is great and has brought positive changes. However, those women who came from distant villages complained and asked us about transportation fees. If we had paid for their transportation costs, it would have motivated them to participate in the training. However, we had budget limitations. (HEW, Shirka woreda, Arsi zone, Oromia)

Selection of WDA leaders

Trainees were recruited by involving different stakeholders: kebele administrators, kebele cabinets, and health professionals from health centers. WDA leaders were primarily selected based on their willingness to contribute to bringing social change for their communities and their experience. Selecting the trainees based on their interest was mentioned as a facilitator:

> Interest to take the training was one of the enrollment criteria, and no one forced them to take part. The ground rule for the training was also prepared by trainees themselves. We had attendance every day for all 13 days. This is one of the reasons that we succeeded in the training with good results. (HEW, Mojana-Wodera Woreda, South Wollo zone, Amhara]

Criteria for the selection of trainees were inconsistent across regions. Views conflicted on choosing WDA leaders based on their educational status. In SNNPR, educational status was not considered as a criterion to select trainees:

> We did not consider educational background while we select them. This is because we believe that knowledge alone without having willingness to share with others or willingness to apply the training into practice is useless. If we get mothers who have little education but are willing to work, we will just select them. (HEW, Meskan woreda, Gurage Zone, SNNPR)

Conversely, a HEP coordinator from Oromia region reported that educational status is one of the criteria during trainee selection:

> We use four basic criteria to select trainees: being a 1-to-30 or 1-to-5 group leader, volunteering to take the training, educational status of completing the first cycle of primary school, and readiness to share their knowledge with other women. [HEP coordinator, Sinana woreda, Bale zone, Oromia]

There were also reports that WDA leaders were excluded from the training if they were pregnant or breastfeeding at the time of the training. As one of the HEP coordinators mentioned,

> During the training, some of the WDA leaders were pregnant, and some of them were old. So, we believed that the pregnant mothers and newly delivered mothers may not be able to come and attend the training. Therefore, we replaced them with their network members. (HEP coordinator, Basona woreda, North Shewa zone, Amhara)

Stakeholder involvement and coordination

Political commitment and coordination among stakeholders played a key role in the success of the CBT program. The program was designed though a collaborative effort by the MoH, the MoE, RHBs, the TVET, funding agencies, and non-governmental implementing partners. WoHOs, health- centers, and health posts also worked together closely to implement the program. A HEP coordinator from Oromia described the collaborative effort of health centers and HEWs by saying,

> We used to get support from health centers. Some healthcare workers were responsible and committed in supporting and monitoring the progress of CBT. They collaborated with HEWs and provided the training. Others were assigned to work at health posts to assist the HEWs. (HEP coordinator, Sinana woreda, Bale zone, Oromia)

However, implementation in some regions and at lower levels was challenged by weak involvement of locally relevant stakeholders from other sectors:

> We began to facilitate the training in collaboration with stakeholders, but we were unable to agree on some issues. For example, there was poor collaboration from the Youth Association Office; we had also little support from the woreda administration. (HEW, Mojana-Wodera Woreda, South Wollo zone, Amhara)

Recognition

Interviewees consistently mentioned that the implementation of CBT lacked consistency with its promises to trainees. A HEP coordinator and trainers mentioned that trainees were informed about the training, assessment, and certification process. Before the training started, the trainees were also informed about the training duration. The absence of assessment and certification resulted in dissatisfaction in the trainees. In some instances, although trainees took the assessment, they were not certified. HEWs mentioned that the delay in certification demotivates trained WDA leaders.

> The CoC assessment was given by other HEWs who came from other places and asked the trainees questions orally. However, they did not get certificates until now. It would be good if the

> certificates were given to maintain their motivation and performance. (HEW, Shirka woreda, Arsi zone, Oromia)

Trainers and HEP coordinators mentioned lack of budget as one of the reasons for delays in assessing and certifying trainees:

> There was no planned budget for assessment and certification. I am sure that there were gaps in providing the certificate. (HEP coordinator, Bati woreda, Oromia special zone, Amhara)

Some trainees also confirmed that certification was delayed even though they took CoC assessments and linked the delay with the COVID-19 pandemic.

> The assessment was very good. But we did not get a certificate because of emergency of the COVID-19 pandemic. Additionally, the evaluation and preparation of reexam for those who failed the exam should be on time. (WDA leader, Mojana-Wodera Woreda, South Wollo zone, Amhara)

On the other hand, in some instances, graduation ceremonies were organized instead of formal assessment and certification. These created a negative impression:

> Initially, we gave a brief orientation about training time, the content, evaluation, CoC processes, and certification. Because of this orientation, trainees were eager to receive their certificates. However, we prepared a graduation ceremony without the assessment and certificates. After that, when we invited them for another activity, they complained that we did not keep our promises. (HEP coordinator, Sinana woreda, Oromia)

Some also reported misunderstandings about who is responsible for giving CoC assessments or who should certify the trainees:

> The woreda informed us [trainers] to evaluate them [trainees]. However, they did not certify them. On our side, we evaluated them both on the theory and on the practice parts. (HEW, Meskan woreda, Gurage zone, SNNPR)

Training Design (Methods)

The comprehensiveness and appropriate design of training materials and the use of different adult teaching–learning methods facilitated the effective implementation of CBT. On the other hand, practical challenges to applying these methods were commonly reported barriers.

Participants acknowledged that the training manual was good in terms of color, font, figures, and text. The pictures and figures included in the manual are self-explanatory, especially for trainees who cannot read and write. They also appreciated the importance of using local languages and considering the levels of literacy of trainees:

> The pictures are understandable. For example, pictures that show how pneumonia is transmitted from one person to the other by drawing the virus spreading through the air while a patient is coughing. Additionally, the training modules were prepared in Afan Oromo; the fonts are bold and visible. These manuals are highly attractive for us. (WDA leader, Tiyo woreda, Bale zone, Oromia)

> These modules were prepared considering the trainees' condition. It is easy to understand and to share it. The language and writing style is clear. (HEW, Bati woreda, Oromia special zone, Amhara)

Despite widely positive views of the current training materials, participants also recommended further improvements. The most frequent recommendation was ensuring adequate availability of existing materials. A HEP coordinator from a WoHO in SNNPR also suggested preparation of the training materials in audio-visual format:

> There was shortage of training manuals. It would also be good to have audio visuals materials. (HEP coordinator, Enemor Ener woreda, Gurage zone, SNNPR)

In addition, trainers mentioned the importance of the exercises and questions presented at the end of some modules to revise and assess the trainees' understanding of each topic. However, some sections do not have such questions:

> Some sections have exercise at the end of the manual and other sections do not. So, we were forced to prepare our own questions to ask trainees to assess their understanding. Those questions that we prepare may not be good enough to summarize the discussed contents. Hence, I recommend to those people who prepared the manual to incorporate questions at the end of each section of the manual. (HEW, Bati woreda, Oromia special zone, Amhara)

Most trainers also reported that the training material is comprehensive and that it covers all important topics. Yet some suggested that the topics should be presented in more detail:

> All-important topics are included in the manual. I recommend expanding the content of the training. They should have more explanation around key concepts without adding new content. (HEW, Sinana woreda, Bale zone, Oromia)

In terms of methods, the training is designed to include both theoretical and practical sessions and exercises at the end of specific topics. Experience-sharing sessions were also incorporated in the implementation of CBT in some places. Trainers and trainees reported that this approach is appropriate and that it made the training both attractive and effective.

> We gave the training in two ways; theoretical and practical. For instance, we gave practical training on how to wash hands, how to clean and keep pure water in the house, how and where to wash housing materials like plates, and how to prepare for solid and wet waste disposal. (HEW, Sinana woreda, Bale zone, Oromia)

Trainers also used innovative approaches in some places, such as the preparation and presentation, by trainees, of dramas and poem on the topics learned and giving reading assignments for participants. Dramas and poems were reported as entertaining while also making trainees understand the topics easily. On the other hand, reading assignments were considered a challenge because of the busy schedules of the women at home.

> For example, trainees prepared drama on what needs to be fulfilled for pregnant women during delivery until she gets an ambulance and how the mother should breastfeed her baby. (WDA leader, Sinana woreda, Bale zone, Oromia)

> Trainers used to give us reading assignments. We usually do not have enough time to read because we have the responsibility of sending our children to school, cooking food for the family, and doing other household work. If I would have time to read the material, I think it would be helpful. (WDA leader, Basona Werana woreda, North Shewa zone, Amhara)

In places where experience sharing is not a practice, participants suggested its inclusion as a teaching method. A WDA leader from Amhara said:

> It would be good if mothers [trainees] could go, see, and learn from other places that have best practices. For example, farmers were previously used to going to other places to learn best practices on crop preparation, irrigation techniques, and other practices. Then, they used to come back to their place and implement the positive things they have observed. Similarly, in our case, these mothers could go and learn from other WDA leaders who have best practices. (WDA leader, Basona Werana woreda, North Shewa zone, Amhara)

Despite the general acceptability and, in some cases, preference for practical sessions, implementation challenges limited the use of practical sessions. For example, absence of a venue and materials limited the implementation of practical sessions.

> To show what was discussed theoretically in practice, it would require having a suitable place. For example, rather than giving the training in an open area under trees' shades, using a training room to show mothers how to prepare food would have been more efficient. This would enable mothers to learn easily, and it would also make them more active in applying their training into practice. (HEW, Meskan woreda, Gurage zone, SNNPR)

> I think the lifesaving is a difficult one. We [HEWs] did not practice it here in the health post; we do not have equipment to practice. We did not train them on the practical aspects, and they are unlikely to practice it. (HEW, Enemor Ener woreda, Gurage zone, SNNPR)

In addition, the nature of some of the interventions was perceived by some of the trainees to be difficult to train through demonstration. Modules on food preparation and environmental sanitation were reported as difficult to cover through demonstration due to the time and other resources they need and the living conditions of the community.

> During the delivery of the hygiene and sanitation module of CBT, you have to practically show what was discussed theoretically in a practical section. However, I found it very challenging. Construction of a latrine as a demonstration exercise and demonstrating handwashing after using the toilet were not easy. (HEW, Meskan Woreda, Gurage zone, SNNPR)

> If possible, food should be prepared in a separate room, which is not feasible in our case because most of the time our mothers live in a single room. (HEW, Meskan woreda, Gurage zone, SNNPR)

In a summary, CBT has been implemented under barriers and facilitators related to issues presented in Table 21.

Table 21. Summary of most frequently reported barriers to and facilitators of CBT, by level

Level	Barriers	Facilitators
Micro-level factors	High illiteracy rate of WDA leaders Beliefs/attitudes contradicting desired behaviors Suboptimal living conditions (*e.g.*, economic hardship and living in a single room) Different levels of understanding among trainees	WDA leaders' altruism Motivated and competent trainers/HEWs Preparedness of trainers Positive working relationship between WDA leaders and HEWs
Meso-level factors	Distance/transportation Social norms, rumors Inappropriate timing (seasons) of training sessions Inconvenient training venues	Existence of functional WDA structures Community acceptance of WDA leaders Flexibility in venue and timing of training Consultation of trainees prior to training
Macro-level factors	Inadequate resource allocation Inadequate stakeholder engagement at lower levels Suboptimal coordination, including lack of clarity on roles and responsibilities of actors Lack of refreshment and incentives during the training Inadequate supervision Inadequacy of training manuals and other materials Low coverage of assessment and certification Inadequate recognition of WDA leaders (below their expectations)	Appropriate design of training manual Adult teaching–learning methods through different approaches incorporating practical sessions Flexibility to incorporate innovative approaches at local levels Selection of trainees based on their interests/volunteerism Flexibility based on the needs of trainees

6 Discussion

6.1 Discussion of key findings

In response to the growing body of evidence and demand for strengthening community engagement, Ethiopia has been working to reposition the role the community can play in improving access to essential healthcare services in the context of primary health care and universal health coverage. Thus, community engagement has become central to the national health strategy in shaping the HEP to promote behavior change, and toward universal and equitable healthcare services coverage over the last two decades (Admasu et al., 2016). As a result, recently, task sharing and the shifting of community-based activities such as health campaigns and community mobilization from HEWs to WDA networks have increased (Admasu et al., 2016; Damtew et al., 2018). This is widely recognized as a cost-effective way to further extend health messages and services to households and communities (Banteyerga, 2011).

However, studies revealed that the WDA approach has been challenged by a lack of proper knowledge and skills to convey the necessary health messages to bring positive changes. The CBT program was

implemented to overcome this challenge by building the capacities of WDA leaders, which is assumed to bring about improvements first in the knowledge and practices of the leaders themselves, then in the functionality of WDA structures, and ultimately in health indicators among network members. Yet limited evidence is available on the effectiveness of CBT and the challenges and facilitators around its implementation. This evaluation provided comprehensive information about the effectiveness of CBT by employing Kirkpatrick's model of training evaluation, which allowed the assessment of trainees' reactions, the effects of the training on trainees' (WDA leaders') knowledge and practices, and the effect of training the leaders on the knowledge and practices of their network members. The evaluation also illuminated some barriers to and facilitators of CBT implementation.

Accordingly, we can draw four principal conclusions from this evaluation: **First**, those who received CBT found it useful and were satisfied with the program. However, heterogeneity exists between and within regions in terms of the selection or recruitment process, the implementation approach, and competency assessment and the duration, quality, and intensity of training. While design-related factors, including content and teaching methods, were drivers of satisfaction, implementation-related issues, including the availability of training materials, convenience of training venues, and suitability of timing of the training sessions, were major sources of dissatisfaction.

The CBT teaching method ranges from giving presentations and demonstrations, to practical sessions involving drama and experience-sharing sessions. WDA leaders favored interactive approaches focusing on practical sessions with hands-on exercise. WDA leaders found training sessions adapted to local culture and delivered in local languages easier to understand. Trainees' complaints included the inconvenience of training venues, poor timing of training sessions, and the absence of recognition or certification.

Second, CBT was found to have a positive association with several aspects of health-related knowledge and behavior, including the use of healthcare services, but effectiveness was not uniform across the ranges of objectives of the training program. Among 30 outcome indicators related to CBT's objectives, CBT was found positively and significantly associated with 13 (43.3%) for WDA leaders and 19 (63.3%) for WDA network members. CBT is not significantly associated with 7 (23.3%) of the expected outcomes for either WDA leaders or network members.

Moreover, significant associations between CBT and health outcomes are variable in their strength. For instance, CBT increased awareness about cervical cancer by 24.3% and participation in WASH campaigns by 16.3%. In contrast, CBT increased the percentage of households covering their food or protecting it from flies by 3.4% and using improved sources of drinking water by 8.6%. We did not observe significant differences among WDA leaders through CBT participation in terms of the use of LARC; PNC coverage; washing hands at critical times; separation of living rooms from the kitchen, animals, and bedrooms; and ownership of improved sanitary facilities.

Patterns in the nature of outcomes that changed and did not change because of CBT largely suggest that CBT has been more effective in disseminating information (*e.g.*, knowledge about FP and HIV, and awareness about cervical cancer and NCDs) and changing behaviors primarily under the control of women (*e.g.*, source of drinking water, participation in WASH campaign, and covering food and protecting food from flies). Practices that require the investment of household resources, beyond the

knowledge and attitude of women (*e.g.*, improved sanitation facilities, separate kitchen, separate sleeping room, and separation of animals from living rooms) tended to remain unchanged because of CBT.

Third, CBT has a spillover effect from WDA leaders to members. The leaders' participation in the training is associated with better HEP-related knowledge and practices among members too. As expected, for most outcomes with positive associations with CBT for both WDA leaders and network members, we have seen a larger effect size on WDA leaders than network members. CBT increased the number of FP methods women known by 2.4, awareness about cervical cancer by 24.4%, and participation in WASH campaigns by 16.3% among WDA leaders, while it increased these indicators by 1.3, 8.2%, and 6.3%, respectively, among network members.

Paradoxically, we have found that CBT was significantly associated with some health outcomes among network members in the absence of significant associations for the WDA leaders, partly due to the smaller sample size for WDA leaders than network members. For seeking treatment for illnesses, washing fruits and vegetables before eating them, and having at least four ANC visits case effect sizes were greater, although not statistically significant, for WDA leaders than for network members.

For the other HEP-related outcomes, including the use of latrines (avoiding open defecation), functional liquid waste disposal pits, adequate water treatment, child vaccination, and FP use, the effect of CBT was higher for WDA network members than for WDA leaders. This could highlight the role of CBT as a mechanism to pass assignments from HEWs to WDA leaders, which may encourage the leaders to promote specific practices although they might not be practicing those behaviors. On the other hand, individuals (WDA network members) may differ not only in their background characteristics but also in how they respond to CBT intervention. Some subjects are likely to benefit most from it and attend CBT (positive selection), while others who had received CBT intervention may be unlikely to benefit most from CBT (negative selection), highlighting the need to revise the selection criteria of WDA leaders and trainees. This issue calls for a more detailed exploration of the mechanisms of influence of WDA leaders in general and that of CBT in particular.

Fourth, the implementation of CBT was affected by factors operating at micro, meso, and macro levels. Factors operating at the micro level are related to the fit between characteristics of trainees and the CBT program, including its processes and the expected roles of trainees. Micro-level challenges identified in our study include WDA leaders' prior knowledge, literacy status, and living conditions, including economic hardship, and differences in the characteristics of WDA leaders attending CBT. These factors are highly related to criteria for the selection of WDA leaders or trainees. Facilitators at this level included altruism among WDA leaders and level of preparedness among trainers.

At the meso level, several community and environmental factors were identified as facilitators of and barriers to CBT, including the convenience of training arrangements and the existence and acceptability of WDA structures. At the macro level, while the design of the training content and methods was considered a facilitator, inadequate resource allocation leading to a shortage of training materials, a lack of refreshment during training sessions, and delayed assessment and certification were key challenges. Suboptimal coordination and a lack of clarity on the roles of different stakeholders, particularly at lower levels, has been a major source of challenges for CBT implementation.

6.2 Strengths and limitations of the study

This evaluation provided comprehensive information about the effectiveness of CBT using the four levels of outcomes from the Kirkpatrick model—reaction, learning, behavior, and results. However, it is not without limitations. The major limitations of the evaluation include (1) nonrandom allocation of WDA leaders to intervention and comparison groups, (2) a lack of data on the CBT process, (3) being underpowered to detect effects on some outcomes, (4) recall and social desirability, and (5) inconsistent data on the sampling frame.

1) Nonrandom allocation of WDA leaders to intervention and comparison groups: The ideal design for studying the effectiveness of CBT would have been a cluster-randomized controlled trial through which kebeles (clusters of WDA leaders and network members) are assigned to receive or not receive CBT through a randomized process. This was impossible because the evaluation began after program implementation. To minimize bias due to the nonrandom distribution of potential confounders, we sampled intervention and comparison groups from similar woredas, which allowed matching by woreda and higher-level confounders, and we used doubly robust estimation techniques to account for uneven distribution of household-level characteristics with potential confounding effects.

2) Lack of data on the CBT process: Primary data on the process of CBT implementation would have allowed objective measurement of implementation intensity, including the level of exposure of every study participant to the intervention. However, this was impossible because of the retrospective nature of our study. We also did not observe how trained WDA leaders transferred the knowledge and skills they obtained from the training. This has limited our ability to explain some of the unexpected findings such as observing effects of CBT on WDA members but not on leaders. This has also limited our ability to understand and account for any possible history of prior or other interventions in the study settings. Moreover, whether WDA leaders in the intervention group completed the training or not was entirely based on HEWs' reports and WDA leaders' personal claims; there was no verification based on objective measures.

3) Underpowered to detect effects on some outcomes: The study was based on a sample of WDA leaders and WDA members for which sample size was determined based on indicators that are relatively sensitive to changes in a short time. These were WASH-related variables. The sample size was adequate to estimate the association between CBT and household-level variables, but outcomes such as child vaccination, coverage of maternal healthcare services, and treatment for sickness were calculated based on a subsample of eligible households. The sample size, particularly the number of WDA leaders, did not yield adequate power to detect effects on these outcomes. Lack of statistical significance should be interpreted carefully.

4) Recall and social desirability: Some responses from WDA leaders were not verified objectively (perceived length of training hours). Practices of some behaviors are also prone to mistakes related to recall and social desirability. However, we do not have a reason to think that these biases differed between the intervention and comparison groups. Therefore, their effect on our effect size estimates safely can be considered negligible.

5) Inconsistent data on sampling frame: The source population (list of potential study woredas) was generated based on a frame provided by RHBs that included a list of woredas and coverage of CBT. However, we found woredas randomly selected based on the obtained frame to be ineligible for the study and replaced seven of eight selected woredas. We tried to mitigate the possible effects of this by selecting replacement woredas from the same zone.

7 Conclusion and Recommendations

7.1 Conclusion

This evaluation is important for policy and public health, as it provides a better understanding of community engagement and valuable insights into capacity-building in the WDA. Our findings showed the following:

- WDA leaders who participated in CBT found it useful. CBT has been satisfactory to participants.
- CBT has brought changes in knowledge and behaviors of WDA leaders in some areas related to WASH, DPC, and maternal and child health. However, the effectiveness of CBT is not uniform across intended program objectives.
- WDA leaders' participation in CBT has improved the functionality of WDA structures. The leaders' participation was also associated with better knowledge and practices in selected health indicators among network members. Nevertheless, the effectiveness of CBT was not uniform across all targeted behaviors.
- The implementation of CBT was facilitated and hindered by factors operating at individual, community, and health-system levels.

7.2 Recommendations

Given the growing momentum and interest in building the capacity of voluntary community health workers, we recommend the following, organized by key findings.

Key finding: Variabilities exist in the implementation of CBT across the regions. WDA leaders may not gain adequate knowledge and skills despite attending CBT.

Recommendation

- Validate whether CBT intervention is being delivered as intended in line with the program model across regions in terms of consistency, adherence, content, process, dose, quality of delivery, and participant responsiveness.
- Monitor and improve the quality of teaching methods, training hours, evaluation mechanisms, and certification, and provide supportive supervision.

Key finding: CBT is effective in improving some of the targeted health outcomes, while it is not for others. This was true for both WDA leaders (direct recipients of the training) and their network members.

Recommendation

- Create diversified approaches customized for each category of targeted health behavior to improving community engagement through building the capacity of community health leaders. No single approach works for all HEP-related outcomes.
- Promote the engagement of men to ensure that women get the support they need to translate their knowledge and skills into behavioral change at the household and community levels.
- Consider CBT part of a broader set of actions to address health problems at the community level rather than a magic bullet to improve health behaviors. The success of CBT requires a conducive environment and enabling conditions and improving the social determinants of health operating at macro and micro levels.
- Conduct further research involving the analysis of cost-effectiveness to see if the results achieved through CBT, particularly those with small effect sizes, are worth the investment needed by the program while continuing the use of this approach to promote the adoption of behaviors about which adequate evidence has been obtained from the current study.
- Monitor implementation fidelity to address implementation-related factors that might have contributed to the lack of or inadequate effect of CBT on some of the targeted health behaviors.

Key finding: Poor engagement of relevant stakeholders affected the implementation of CBT.

Recommendation

- Establish strategic partnerships and promote coordination among key stakeholders (such as the Ministry of Women and Children's Affairs, agricultural extension services, and the WASH Committee). This will lead to the creation of synergy to address the identified implementation gaps and other social determinants of WASH, FP, and maternal and child healthcare use and the continued adoption of an evidence-based approach for the capacity development of WDA leaders.

Key finding: Inadequate resource allocation for CBT has been a major source of several challenges the program faced at both kebele and woreda levels.

Recommendation

- Ensure sustainable capacity development for WDA leaders by balancing the speed of scaleup with the allocation of adequate resources to make sure that all the required processes, including preparation, implementation, and follow-up support, are in place.

Key finding: Even for outcomes where CBT was found effective, adoption of HEP-related behaviors was suboptimal among WDA leaders and their network members.

Recommendation

- If WDA networks are identified to be a viable approach, ensure sustainable continuous capacity development for WDA leaders involving regular supportive supervision, conduct periodic program performance assessments, and provide refresher training. These can be reinforced by designing appropriate incentive mechanisms to keep WDA leaders motivated.

- Continuously monitor CBT based on key performance indicators for WDA leaders, including indicators for assessing reaction, learning, behavior, and influence on network members. Support monitoring with follow-up and improvement actions on the quality of implementation, including teaching methods, training hours, assessment mechanisms, and certification.

References

ADMASU, K., BALCHA, T. & GHEBREYESUS, T. A. 2016. Pro–poor pathway towards universal health coverage: lessons from Ethiopia. *Journal of Global Health*, 6.

ASSEFA, Y., GELAW, Y. A., HILL, P. S., TAYE, B. W. & VAN DAMME, W. 2019. Community health extension program of Ethiopia, 2003–2018: successes and challenges toward universal coverage for primary healthcare services. *Globalization and Health,* 15**,** 24.

BANTEYERGA, H. 2011. Ethiopia's health extension program: improving health through community involvement. *MEDICC Rev,* 13**,** 46–9.

BATES, R. J. E. & PLANNING, P. 2004. A critical analysis of evaluation practice: the Kirkpatrick model and the principle of beneficence. *Evaluation and Program Planning*, 27**,** 341–347.

BILAL, N. 2012. Health Extension Program: An innovative solution to public health challenges of Ethiopia - a case study. *Health Finance and Governance***,** 12.

BILAL, N., HERBST, C., ZHAO, F., SOUCAT, A. & LEMIÈRE, C. 2011. Health extension workers in Ethiopia: improved access and coverage for the rural poor. *Yes Africa can success stories from a dyn cont*.

BITTON, A., FIFIELD, J., RATCLIFFE, H., KARLAGE, A., WANG, H., VEILLARD, J. H., SCHWARZ, D. & HIRSCHHORN, L. R. 2019. Primary healthcare system performance in low-income and middle-income countries: a scoping review of the evidence from 2010 to 2017. *BMJ Global Health,* 4**,** e001551.

DAMTEW, Z. A., KARIM, A. M., CHEKAGN, C. T., FESSEHA ZEMICHAEL, N., YIHUN, B., WILLEY, B. A. & BETEMARIAM, W. 2018. Correlates of the Women's Development Army strategy implementation strength with household reproductive, maternal, newborn and child healthcare practices: a cross-sectional study in four regions of Ethiopia. *BMC Pregnancy Childbirth,* 18**,** 373.

FMoH 2008. Guideline for Implementation of Health Extension Program. Addis Ababa, Ethiopia.

FUNK, M. J., WESTREICH, D., WIESEN, C., STURMER, T., BROOKHART, M. A. & DAVIDIAN, M. 2011. Doubly robust estimation of causal effects. *Am J Epidemiol,* 173**,** 761–7.

KRISHNAN, P. 2019. A review of the non-equivalent control group post-test-only design. *Nurse Res,* 26**,** 37-40.

LA DUKE, P. J. P. S. 2017. How to evaluate training: using the kirkpatrick model. 62**,** 20.

MAES, K., CLOSSER, S., TESFAYE, Y., GILBERT, Y. & ABESHA, R. 2018. Volunteers in Ethiopia's Women's Development Army are more deprived and distressed than their neighbors: cross-sectional survey data from rural Ethiopia. *BMC Public Health,* 18**,** 258.

MERQ CONSULTANCY PLC 2019. National assessment of the Health Extension Program in Ethiopia: abridged report.

TEKLU, A. M., ALEMAYEHU, Y. K., MEDHIN, G., FENTAYE, F. W., DADI, T. L., TSEHAY, Y. E., HILAWE, E. H., GOBEZIE, W. A., CHERE, G. A., TESFAY, T. G., GEBREMEDHIN, M., TEKLE, M. G., ASSEFA, D. T., AMIN, H. M., GEZIE, A. M., SENTURIA, K., CHALLA, M. D., FIDA, A. T., TADESSE, T., HAILESELASSIE, W., MULUGETA, T., DINBERU, B., GEREMEW, E., JEBENA, M. G., ESHETE, M. T. & MEDHANIYIE, A. A. 2020. National assessment of the Ethiopian Health Extension Program 2019. MERQ Consultancy PLC.

YITBAREK, K., ABRAHAM, G. & MORANKAR, S. 2019. Contribution of Women's Development Army to maternal and child health in Ethiopia: a systematic review of evidence. *BMJ Open,* 9**,** e025937.

List of Annexes

Annex 1: Listing and sampling formats

Annex 2: Data collection tools

Annex 3: List of study woredas: Initial sample and replaced woredas