

Effect of Augmented Capacity Development Interventions (ACDI) on the performance of data quality in the Routine Health Information System (RHIS) among health workers in public health institutions of Gofa Zone, Southern Ethiopia: a cluster randomized controlled trial

Journal:	<i>SAGE Open Medicine</i>
Manuscript ID	SOM-25-0809
Manuscript Type:	Original Research Article
Date Submitted by the Author:	02-Jul-2025
Complete List of Authors:	Doka, Bedilu; Hawassa University College of Medicine and Health Sciences, School of Public Health; Regional State Health Bureau , Health and Health Related Product Quality and Services Regulatory Authority Worku, Abebaw; Ethiopia Ministry of Health, John Snow Inc Research and Training Institute, Inc., Ethiopia Data Use Partnership Negeri, Keneni; Hawassa University College of Medicine and Health Sciences, School of Public Health Kassa, Dejene; Hawassa University College of Medicine and Health Sciences, Public Health
Keywords:	Routine Health Information System, Health data quality, Augmented Capacity Development Interventions
Additional Keywords:	Health workers, Public health institutions, Southern Ethiopia, Cluster randomized controlled trial
Abstract:	<p>Objectives: Strengthening data quality in the routine health information system is vital for the performance of health service outcomes. However, implementation of the routine interventions to improve data quality in the existing health system has been found poor in Ethiopia. This study was aimed to examine the effect of Augmented Capacity Development Interventions (ACDI) on the performance of data quality in the routine health information system.</p> <p>Methods: A two arm, parallel group, cluster-randomized control trial was implemented from July 1, 2023 to February 29, 2024 in public health institutions of Gofa Zone, Southern Ethiopia.. A total of 72 clusters of health institutions and 304 health workers (154 in the intervention group and 150 in the control group) were included in the study. The implemented interventions include training, supportive supervision, mentorship, and recognition. General Linear Mixed Model was applied for analysis.</p> <p>Results: The data quality practice has significantly improved after the implementation of the ACDI packages ($\beta = 0.17$; 95% CI: 0.05, 0.30; $p = 0.007$). Encouragement ($\beta = 0.53$; (95% CI: 0.29, 0.77; $p < 0.001$), ease of data management ($\beta = 0.14$; 95% CI=0.07, 0.22, $p < 0.001$), information use ($\beta = 0.15$; 95% CI: 0.08, 0.23), $p < 0.001$), availability of written guideline ($\beta = 0.14$; 95% CI: 0.04, 0.24, $p = 0.006$), the combined effects of encouragement and training ($\beta = 0.44$; 95% CI: 0.23, 0.65; $p = 0.010$), and consistence use of tools ($\beta = 11$; 95% CI: 0.02, 0.21; $p = 0.023$ were significant predictors of the change in the data quality.</p>

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

	<p>Conclusion: The ACDI packages implemented in this study effectively influenced data quality improvement. Key predictors of data quality practices included an encouraging system, ease of data management, written guidelines, supportive supervision, and training. Therefore, the interventions are recommended to be scaled up.</p> <p>Trial registration: ID: PACTR202212472091194, registered on 14 December 2022.</p>



Effect of Augmented Capacity Development Interventions (ACDI) on the performance of data quality in the Routine Health Information System (RHIS) among health workers in public health institutions of Gofa Zone, Southern Ethiopia: a cluster randomized controlled trial

Bedilu Kucho Doka^{1, 2*}, Abebaw Gebeyehu Worku³, Keneni Gutema Negeri¹, Dejene Hailu Kassa¹

¹School of Public Health, College of Medicine and Health Sciences, Hawassa University, Hawassa, Sidama, Ethiopia

²Health and Health Related Product Quality and Services Regulatory Authority, South Ethiopia Regional State Health Bureau, Jinka, Ethiopia

³John Snow Inc Research and Training Institute, Inc., Ethiopia Data Use Partnership, Ministry of Health, Addis Ababa, Ethiopia

*Corresponding author: bedilukucho54@gmail.com

Abstract

Background: Strengthening data quality in the routine health information system is vital for the performance of health service outcomes. However, implementation of the routine interventions to improve data quality in the existing health system has been found poor in Ethiopia. This study was aimed to examine the effect of Augmented Capacity Development Interventions (ACDI) on the performance of data quality in the routine health information system.

Methods: A two arm, parallel group, cluster-randomized control trial was implemented from July 1, 2023 to February 29, 2024 in public health institutions of Gofa Zone, Southern Ethiopia.. A total of 72 clusters of health institutions and 304 health workers (154 in the intervention group and 150 in the control group) were included in the study. The implemented interventions include training, supportive supervision, mentorship, and recognition. General Linear Mixed Model was applied for analysis.

Results: The data quality practice has significantly improved after the implementation of the ACDI packages ($\beta = 0.17$; 95% CI: 0.05, 0.30; $p = 0.007$). Encouragement ($\beta = 0.53$; (95% CI: 0.29, 0.77; $p < 0.001$), ease of data management ($\beta = 0.14$; 95% CI=0.07, 0.22, $p < 0.001$), information use ($\beta = 0.15$; 95% CI: 0.08, 0.23), $p < 0.001$), availability of written guideline ($\beta = 0.14$; 95% CI: 0.04, 0.24, $p = 0.006$), the combined effects of encouragement and training ($\beta = 0.44$; 95% CI: 0.23, 0.65; $p = 0.010$), and consistence use of tools ($\beta = 11$; 95% CI: 0.02, 0.21; $p = 0.023$ were significant predictors of the change in the data quality.

Conclusion: The ACDI packages implemented in this study effectively influenced data quality improvement. Key predictors of data quality practices included an encouraging system, ease of data management, written guidelines, supportive supervision, and training. Therefore, the interventions are recommended to be scaled up.

Trial registration: ID: PACTR202212472091194, registered on 14 December 2022.

Keywords: Augmented Capacity Development Interventions, Data quality, Routine Health Information System, Health workers, Public health institutions, Southern Ethiopia, Cluster randomized controlled trial

For Peer Review

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1. Introduction

Routine health information system (RHIS) is the process of collection, interpretation, utilization, and dissemination of routine health data targeted to improve health system performance and health service outcomes (1). Data quality in health is a multifaceted concept commonly expressed in the dimensions of data accuracy, completeness and timeliness to generate reliable information for decision making process (2). The ultimate purpose of a RHIS is to produce quality information to be presented for an evidence-based decision making process (3).

In low- and middle-income countries, as revealed by scientific evidence, the overall health data quality was much below the expected national standard (4–6). In these countries, the performance of routine health data is challenged by poor data management skills, lack of commitment from managing bodies, shortage of monitoring and evaluation system, inadequate infrastructure, and shortage and high turnover of skilled staff (7–9). Therefore, the performance of existing health system interventions on RHIS has been found to be poor in the developing world (10).

Most studies in developing nations reported that the rate of routine data quality was far below the World Health Organization (WHO) standard of 90% (10–12). The completeness rate of District Health Information Software, version 2 (DHIS2) data was only 60% whereas under-reporting ranges between 10% – 60% according to a study in Nigeria (10). Data quality assessment in Mali indicated that the rate of data accuracy at the health facility level was 45% and timeliness of 27%. Data accuracy was 68% at the regional and 54% at the central levels (12).

A survey in the Tigray region of Ethiopia revealed that the data completeness rate of registration books and reports were 54% and 56%, respectively. Similarly, an internal consistency

of the data was 39% (11). A finding from Oromia region of Ethiopia showed that timeliness was 70%, completeness was 86%, and accuracy was 48% (13). Another similar study carried out in Dire Dawa city of Ethiopia indicated that the level of data quality was 75% in health institutions (14). A study carried out in Southern Ethiopia indicated Antenatal Care (ANC) four visit, postnatal care and fully immunized were over reported as 24%, 21% and 16% respectively (15).

There is no individual intervention that could be operated separately to enhance the quality of data in RHIS, but the implementation of the combination of interventions is recommended to assure high quality of health data (16). A survey in Pakistan revealed that data accuracy has increased from 40% to 75% after implementation of DHIS platform as intervention (17,18). A study in Nigeria revealed improvements in data quality metrics after implementation of interventions. Accordingly, report completeness rate has improved from 72% to 82%, timeliness increased from 60% to 72%, the report content completeness increased from 62% to 68% (10).

A study in Northwest Ethiopia reported that data consistency has improved from 84.0% to 99.5%, data recording completeness from 69% to 96%, and that of report timeliness increased from 66% to 100% after implementation of intervention packages comprising training, supportive supervision and feedback provision (19). However, the study addressed only one primary hospital of the area where its representativeness is very limited.

Limited previous intervention studies were available on data quality in RHIS. Even though few studies were conducted previously, the interventions implemented were not comprehensive, with many focusing on just a few packages, such as training and guideline provision. These studies were done only on a single or very few facilities or districts; therefore, their representativeness is questionable. Still other studies were design-related limitations such as a lack of control groups and inability of considering variation of the nature of the outcome among health institutions.

Therefore, this study aims to evaluate the effect of Augmented Capacity Development Interventions (ACDI) on the performance of data quality in the RHIS among health workers in public health institutions of Gofa Zone, Southern Ethiopia region.

2. Methods

2.1. Study Setting

The study was carried out in health institutions of Gofa Zone, Southe Ethiopia region. According to the central statistical agency of Ethiopia, the current population of the Zone is approximately 713,854. Among these, 357,359 (50.1%) are men, and 356,495 (49.9%) are women, with a total of 145,684 households. A total of 1510 health professionals are deployed in 11 districts, 26 health centers, 179 health posts, 2 governmental hospitals.

2.2. Trial design

A two arm, parallel group, cluster randomized controlled trial design was adopted. This design was selected in order to minimize experimental contamination between groups as the intervention is implemented at group level. The baseline data were collected from April 1 to 30, 2023. The intervention was implemented from July 1, 2023 to February 29, 2024. The end-line data were collected from April 1 to 30, 2024.

2.3. The participants

The source population comprised all districts, public health facilities, and health workers existing in the zone. Randomly selected health institutions and health workers constituted the study population.

2.4. Eligibility

Inclusion criteria

All health workers, including those serving in different departments, health posts, and heads of health institutions were included in the study. Administrative district health offices and functional public health facilities, including hospitals, health centers, and health posts were included.

Exclusion criteria

Newly established (2 health posts), nonfunctional (4 health posts) and privately owned health facilities were not considered in this study. The health workers who were not available during baseline data collection (5 health workers); who intended to leave the institution within eight months immediately prior to the baseline data collection (6 health workers); and who did not receive the intervention or dropped out at some point (13 health workers) were also excluded.

2.5. Sample size determination

The study applied the assumptions of confidence level of 95%, marginal error of 5%, and intervention to control ratio of 1:1 to determine the sample size. The sample size was calculated by considering the percent of data quality in comparison group of 33% (20). Power of 90% was assumed to detect 30% difference in rates between the two groups. Since the study was a cluster design, ICC of 0.35 and average cluster size of 4.3 were utilized from previous related study (21). The design effect of 2.2 and non-response rate of 10% was considered. Therefore, a total of 309 health workers of both groups were targeted to be recruited from 72 health institutions including 6 districts, 2 hospitals, 18 health centers and 46 health posts. However, in the baseline, 5 respondents were non-respondents, and 13 were lost to follow-up in the endpoint data collection.

2.6. Sampling procedures and randomization

A multistage stratified cluster sampling technique was employed to select study institutions. The zone has 11 districts (seven rural and four urban). First, the rural-urban stratification of the districts

was implemented. Then, four from seven rural districts (Demba Gofa, Zalla, Gezegofa and Melokoza) and two (Sawla and Laha) from four urban districts of the zone were selected by simple random sampling technique. The districts were selected with all their respective health facilities. Based on this, 18 health centers and 2 hospitals from selected districts were considered for the study. Additionally, a total of 46 health posts were proportionally allocated from each of the corresponding health centers. Regarding the groups, two randomly selected rural districts, Geze Gofa and Demba Gofa, as well as one urban district (Sawula) with all their respective facilities were included under intervention group. Therefore, a total of 37 health institutions were included under the intervention institutions that constituted 24 health posts, 9 health centers, 1 hospital, 2 rural and 1 urban district. On the other hand, two randomly selected rural districts (Zalla and Mello Koza) and one urban district (Laha) with their respective health facilities including 22 health posts, 9 health centers and 1 hospital were the part of overall 35 control institutions. It is based on the intervention to control ratio of 1:1. Regarding the selection of health workers, the heads of the health institutions and randomly selected participants from the Outpatient Department (OPD), Maternal and Child Health (MCH), emergency, dispensary, laboratory, Health Management Information System (HMIS) departments and office management were included. Health workers were recruited at baseline, before the randomization of clusters into groups was carried out.

Allocation sequence, concealment and blinding

Sequence generation

Before the implementation of the randomization process, districts were first stratified by location type (i.e., urban or rural). Then, to reduce the risk of experimental contamination, districts were allocated using the block randomization procedure. Three (one urban and two rural) adjacent and contiguous districts were grouped into one block and the other three (one urban and two rural)

districts were sorted in to the other block. Finally, the blocks were randomly selected and allocated into either intervention or control groups. The two blocks of the districts were separated geographical buffers of unselected districts, special zones and rivers. Although there is still some territorial connection between certain control and intervention districts, the risk of contamination is not significant, as there is no physical proximity between the facilities because a buffer zone with a minimum distance of ten kilometers was established.

Allocation concealment

In order to minimize the selection bias and ensure unpredictability, the assignment of the blocks to the either arm has been done by an independent researcher from Arbaminch University of Ethiopia, who was unaware of the study group assignments, applied sealed envelopes for the group allocation.

Blinding

In order to avoid any bias on study results, the outcome assessors were withheld about the interventions provided as they were deployed from unselected districts. The blinding of program implementers and study participants was not possible as they provide and receive the open-level trial. However, the control groups were kept unaware of what the intervention groups received.

2.7. Variables

Outcome variables

Data quality is a multidimensional construct expressed in terms of timeliness, accuracy, and completeness. **Data accuracy rate** was determined by comparing the data recorded in the summary forms (HMIS report) with the data in the registers from health facilities. For the districts, the results of aggregated reports were compared with recounted reports of indicators from each reporting facilities (22). A sample of 13 indicators including ANC first, pentavalent third dose, postnatal care,

contraceptive acceptance rate and malaria testing rate were used for assessing data accuracy. If the value of the verified data exceeds 100%, which is considered as under-reporting whereas less than 100% will be over-reporting (23). A margin of error will be considered acceptable at 5% (95% to 105%), fair at 10% (90% to 110%), unacceptable at greater than 10% (below 90% and above 110%) (24)

Completeness was explained by report completeness and completeness of indicator data. **Report completeness** was measured by the number of institutions reported against total institutions expected for a specified period (13). Report completeness is computed only for districts and health centers, but not for health posts and hospitals, as the latter do not have independent reporting facilities under them. **Completeness of indicator data** was examined by all the relevant data elements in a reports were filled and measured by calculating the proportion of data cells filled for all cases from the total expected in the reports (25). To identify data completeness, the aforementioned 13 data elements were computed (26).

Timeliness of reporting was determined by analyzing health institution summary reports that were remitted to the next level within a predetermined reporting period based on the Ethiopian national reporting schedule (27). Report timeliness was also calculated only for districts and health centers.

Data quality practice

Data quality practice is the primary outcome in this study. It is a composite construct measured by the level of agreement on 11 items using a Likert scale format, where a score of 5 represents 'Strongly Agree,' 4 represents 'Agree,' 3 represents 'Neutral,' 2 represents 'Disagree,' and 1 represents 'Strongly Disagree.' In this case, the perception of respondents about how their institution performs in terms of data quality is assessed.

The items used to measure this construct include: provision of quality healthcare and generation of sufficient data, quality of documentation, accuracy in data compilation, effectiveness of data communication; efficiency of data collection, sharing, and reporting systems; appropriateness and accuracy of data for quality decision-making, timeliness of reporting, assurance of completeness of information in records and reports, conducting data quality reviews before reporting, and promotion of integrity in data management.

The 11 item scores for each respondent are averaged to produce a single value, and finally, the mean of these scores is computed, treating the variable as continuous.

Predictors like perceived ease of data management, level of information utilization, supervision quality and perceived level of job satisfaction were also computed using the same procedure as data quality practice.

The Intervention

The study implemented the ACDI packages that were targeted to improve data quality in RHIS among selected health institutions. The following interventions were implemented to improve the knowledge, practice, and skill of the experts and the system to produce quality data.

Training

Training process was led by intervention team members, who had received initial master training and were assigned to deliver the intervention. The training of the intervention community members, selected staff from different departments, comprising HMIS officers, managers, Performance Monitoring Members (PMT) members, and health extension workers was carried out. The training was organized in their respective institutions. At the beginning of the intervention process, a four-day initial training for the intervention community was organized on six relevant modules. Then

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

assessment-based on-the-job training workshops were conducted in every two months by the intervention team.

Supportive supervision and feedback

In this study, checklist based supervisory assessment and action oriented written feedback system on data management process were implemented for one day per every two month period. A total of four supportive supervision visits were delivered for each intervention institution during the eight month implementation period.

Mentorship

The mentorship program was organized in every three months each for two days. Trained experts, from the members of the intervention team, were the mentors and the intervention communities were the mentees. Close observation, guidance and spot assistance based on initial and subsequent assessment findings were addressed and forwarded for discussion.

Monitoring and Evaluation

In this project, every two-month review meetings and learning workshops based supervision findings were organized following the second day of the supervisory assessments. Monitoring was implemented in an ongoing process by the institution representatives for the correction and implementation of supervision feedbacks as part of self-assessment.

Recognition

This project involved individual and organizational certification of outstanding performance, appreciating the workers, departments and institutions by using verbal communication, encouraging the workers to share their successful experience for other institutions, promoting and scaling up

their effective intervention approaches. The recognition process was practiced along with learning workshops.

We state that the intervention is augmented, as substantial modifications were made to conventional routine practices. Overall, the intervention institutions have received all the stated interventions while the control institutions were given the training modules at the end of the study period. Different Standard Operating Procedures (SOPs) were established for the implementation of each intervention packages and corresponding activities.

2.8. Data collection procedures

A total of eight data collectors and three supervisors were deployed for data collection after three-day intensive training. The questionnaire was designed in English version and translated to the Amharic language for better understanding of respondents. The data were collected with a structured, pre-tested, and standardized questionnaire customized from Performance of Routine Information System Management (PRISM) assessment tools (28). The data were collected using face to face interview questionnaires, document review templates, physical observation checklist that were also used for organizing surveys, reviewing documents and conducting observations. An electronic data collection process has been implemented using the Kobo Toolbox.

2.9. Quality control

Data collectors and supervisors have received an intensive training on data collection protocol. We have adapted a standard Performance of Routine Information System Management (PRISM) assessment tools (28) in designing the questionnaire for the study. During the manuscript writing process, we have thoroughly followed the guideline of the Consolidated Standards of Reporting Trials (CONSORT) with Extension to Cluster Randomized Trials to assure a standard reporting process of the trial.

2.10. Statistical analysis

Data were exported to Stata 17 for analysis. Descriptive statistics including frequencies, proportions, mean, and standard deviation were computed. The 95% confidence interval was used. Repeated measures analysis was conducted using a General Linear Mixed Effects Model regression. Variables with a p-value of less than 0.25 in the bi-variable analysis were entered into the multivariable regression analysis. A p-value of less than 0.05 in the multivariable regression analysis were reported to identify predictor variables significantly associated with the outcome variable (29). We applied the intention-to-treat (ITT) approach to account for missing values from 13 participants lost to follow-up, in order to preserve the benefit of randomization and prevent bias caused by dropouts.

Results

Participant flow

During the baseline data collection, a total of 304 participants (154 intervention and 150 control) were surveyed from 71 health institutions (37 intervention and 35 control). However, for the end-line data collection, a total of 291 health workers participated, resulting in a loss to follow-up of 13 (4.2%) participants. On the other hand, 70 health institutions were included in the final data collection, as two health extension workers from two health posts were lost to follow-up (Figure 1).

Characteristics of the study participants

About half (50.5%) of the participants who completed the follow up were from the intervention institutions. Among the 291 study participants, 184 (63.2%) were males and 227 (78.0%) were from rural health institutions. The average age of the participants was 30.05 years (SD = 3.81), while the median work experience was 6.16 years (SD = 3.70) (Table 1).

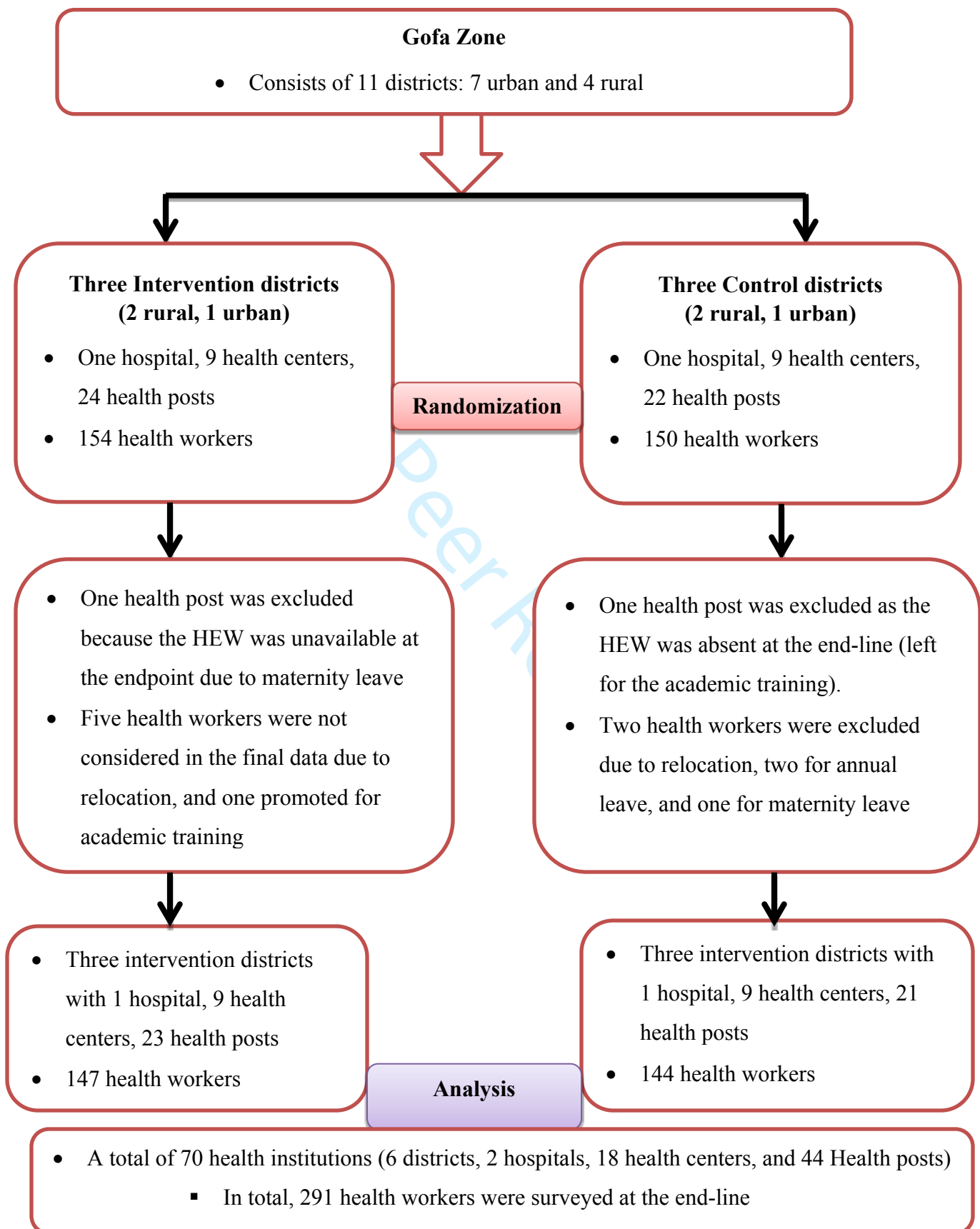


Figure 1. Flow diagram illustrating the selection procedure for health workers in public health institutions of Gofa Zone, Southern Ethiopia, 2024.

Table 1. Characteristics of study participants among health institutions of Gofa Zone, South Ethiopia Region, 2024.

Characteristic	Baseline			End-line		
	Intervention Group (n=154)	Control Group (n=150)	p-value (adjusted)	Intervention Group (n=147)	Control Group (n=144)	p-value (adjusted)
Age (years), Mean ± SD	29.25±3.57	30.91±3.80	0.168	29.17±3.58	30.96± 3.85	0.151
Sex, n (%)						
Male	99 (64.3)	95 (63.3)	0.988	93 (63.3)	91 (63.2)	0.964
Female	55 (35.7)	55 (36.7)		54 (36.7)	53 (36.8)	
Educational status, n (%)						
Certificate and lower	3 (1.9)	0 (0.0)	0.994	3 (2.7)	0 (0.0)	0.994
Diploma	88 (57.1)	110 (73.3)	0.801	85 (57.8)	105 (72.9)	0.794
Degree	59 (38.3)	38 (25.3)	0.933	55 (37.4)	37 (25.7)	0.920
Masters and above	4 (2.6)	2 (1.3)		4 (2.7)	2 (1.4)	
Department of the participant, n (%)						
Office management	26 (16.9)	28 (18.7)	0.882	24(16.3)	26 (18.1)	0.877
MCH	39 (25.3)	47 (31.3)	0.824	36 (24.3)	45 (31.3)	0.807
HMIS unit	11 (7.1)	12 (8.0)	0.906	12 (8.2)	12 (8.3)	0.931
OPD	18 (11.7)	15 (10.0)	0.992	17 (11.6)	15 (10.4)	0.962
Emergency	15 (9.7)	7 (4.7)	0.937	15 (10.2)	7 (4.9)	0.831
Laboratory	11 (7.1)	7 (4.7)	0.944	11 (7.2)	6 (4.2)	0.896

1							
2							
3	Dispensary	10 (6.5)	12 (8.0)	0.880	9 (6.1)	12 (8.3)	0.844
4	Health post	24 (15.6)	22 (14.7)		23 (15.6)	21 (14.6)	
5							
6	Work experience (years), Median \pm SD	6.16 \pm 3.59	6.33 \pm 3.80	0.853	6.16 \pm 3.56	6.33 \pm 3.84	0.801
7							
8	Conducting DQA, n (%)						
9							
10	Yes	72 (46.8)	23 (15.3)	0.067	128 (87.1)	85 (59.0)	0.087
11	No	82 (53.2)	127 (84.7)		19 (26.8)	59 (41.0)	
12							
13	Written guideline, n (%)						
14							
15	Yes	53 (34.4)	49 (32.7)	0.925	129 (87.8)	45 (31.3)	0.001
16	No	101 (65.6)	101 (67.3)		18 (22.2)	99 (68.7)	
17							
18	Strategic plans, n (%)				130 (88.4)	50(34.7)	<0.001
19							
20	Yes	89 (57.8)	49 (32.7)	0.291	17 (11.6)	94(65.3)	
21	No	65 (42.2)	101 (67.3)				
22							
23	PMT availability, n (%)						
24							
25	Yes	125 (81.2)	122 (81.3)	0.836	137 (93.2)	121 (84.0)	0.204
26	No	29 (18.8)	28 (18.7)		10 (6.8)	23 (16.0)	
27							
28	Training on RHIS, n (%)						
29							
30	Yes	39 (25.3)	35 (23.3)	0.873	128 (87.1)	101 (70.1)	0.385
31	No	115 (74.7)	115 (76.7)		19 (12.9)	43 (29.9)	
32							
33	Receive feedback on RHIS, n (%)						
34							
35	Yes	60 (39.0)	59 (39.3)	0.567	136 (92.5)	108 (75.0)	0.365
36	No	94 (61.0)	91 (60.7)		11 (7.5)	36 (25.0)	
37							
38	Receive supportive supervision on						
39	RHIS, n (%)						
40							
41	Yes	26 (16.9)	16 (10.7)	0.845	136 (92.5)	107 (74.3)	0.244

Consistently use standard tools, n (%)	No	128 (83.1)	134 (89.3)		11 (7.5)	37 (15.7)	
	Yes	67 (43.5)	81 (54.0)	0.353	137 (93.2)	93 (64.6)	0.031
Supervision quality, Mean ± SD	No	87 (56.5)	69 (46.0)		10 (6.8)	51 (35.4)	
Availability of rewarding system, n (%)		2.35±0.28	2.42±0.18	0.510	3.82±0.59	2.33±0.43	<0.001
	Yes	14 (9.1)	20 (13.3)	0.925	130 (88.4)	19 (13.2)	<0.001
Encouraging system for good performance, n (%)	No	140 (90.9)	130 (86.7)		17 (11.6)	125 (86.8)	
	Yes	9 (5.8)	8 (5.3)	0.953	142 (96.6)	8 (5.6)	<0.001
Perceived level of your job satisfaction on RHIS, Mean ± SD	No	145 (94.2)	142 (94.7)		5 (3.4)	136 (94.4)	
Ease of data management, Mean ± SD		2.42±0.25	2.46±0.17	0.484	3.84 ±0.58	2.41±0.58	<0.001
Information utilization, Mean ± SD		2.29±0.32	2.31±0.32	0.529	2.95±0.78	2.33±0.73	0.01
		2.319±0.171	2.275±0.126	0.319	3.023±0.908	2.263±0.606	0.011

MCH = Maternal and Child Health, HMIS unit = Health Management Information System unit, OPD = Outpatient Department, PMT = Performance Monitoring Team, SD = Standard Deviation, DQA = Data Quality Assessment, RHIS = Routine Health Information System

Data quality dimensions

Report completeness and report timeliness

At the beginning of the study, the average report completeness of the health institutions was 96.20% (95% CI: 93.40, 97.50; $p = 0.065$), with the intervention institutions at 98.70% and the control institutions at 93.60%. However, by the end of the study, no significant change was observed, with the report completeness slightly dropping to 92.70% (95% CI: 85.20, 100.20; $p = 0.146$), with the intervention institutions at 98.10% and the control institutions at 87.3%. The overall average report completeness rate was 95.80% (95% CI: 93.02, 98.62; $p = 0.054$).

Regarding report timeliness, the result improved from 53.30% (95% CI: 34.10, 72.60; $p = 0.627$) at baseline to 76.50% (95% CI: 63.20, 89.50; $p = 0.003$) at end-line, with an overall average timeliness score of 78.20% (95% CI: 64.90, 91.50; $p = 0.001$) (Figure 2).

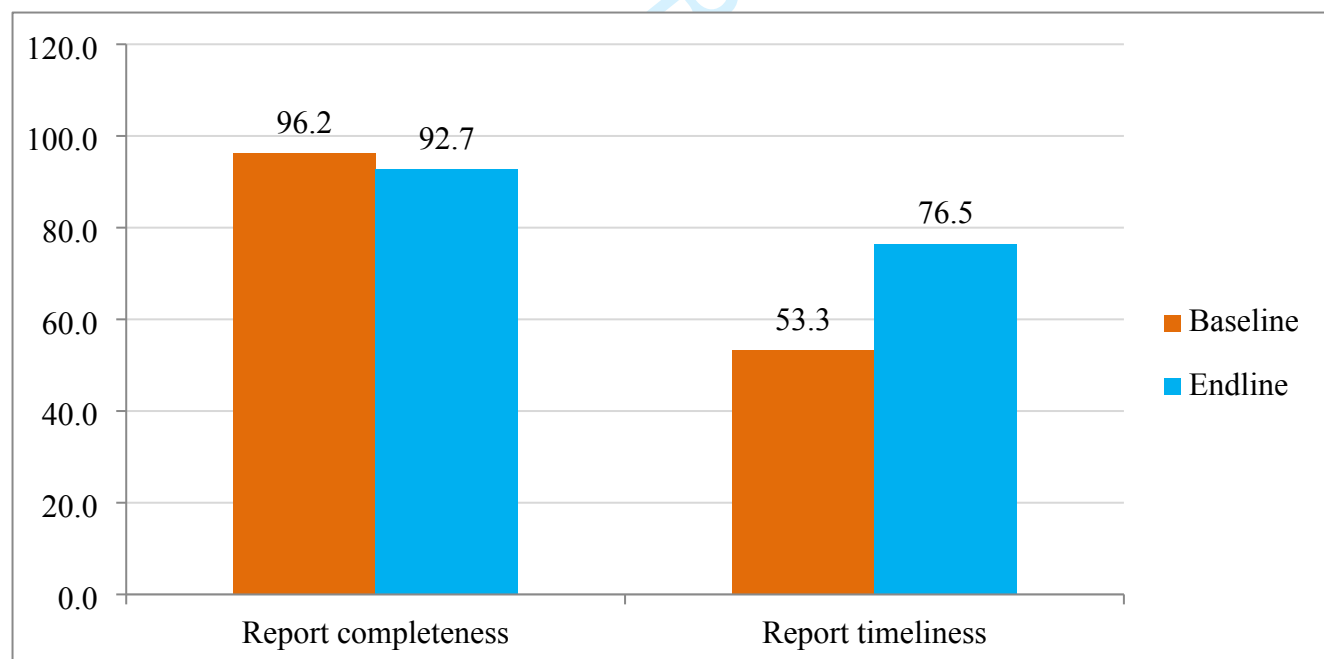


Figure 2. Report completeness and timeliness among health institutions in Gofa Zone, Southern Ethiopia, 2024 (n = 24).

Data accuracy and data completeness

Data completeness in the health institutions showed a significant change from a baseline score of 93.51% (95% CI: 90.12, 96.88; $p = 0.936$) to an end-line score of 93.44% (95% CI: 90.54, 96.33; $p < 0.001$). The overall data completeness score was 93.36% (95% CI: 91.02, 95.70; $p < 0.001$).

Within this context, the intervention institutions had an average score of 99.13%, while the control institutions scored 87.59%.

On the other hand, the data accuracy score increased from 89.40% (95% CI: 85.32, 93.46; $p = 0.895$) at baseline to 95.63% (95% CI: 92.82, 98.45; $p < 0.001$) at the end point of the study, indicating a significant change over time. However, there was no statistically significant difference between the groups, with an overall average data accuracy rate of 94.65% (95% CI: 92.30, 97.00; $p = 0.087$). In this case, the average data accuracy score in the intervention institutions was 99.69%, compared to 92.60% in the control group (Figure 3).

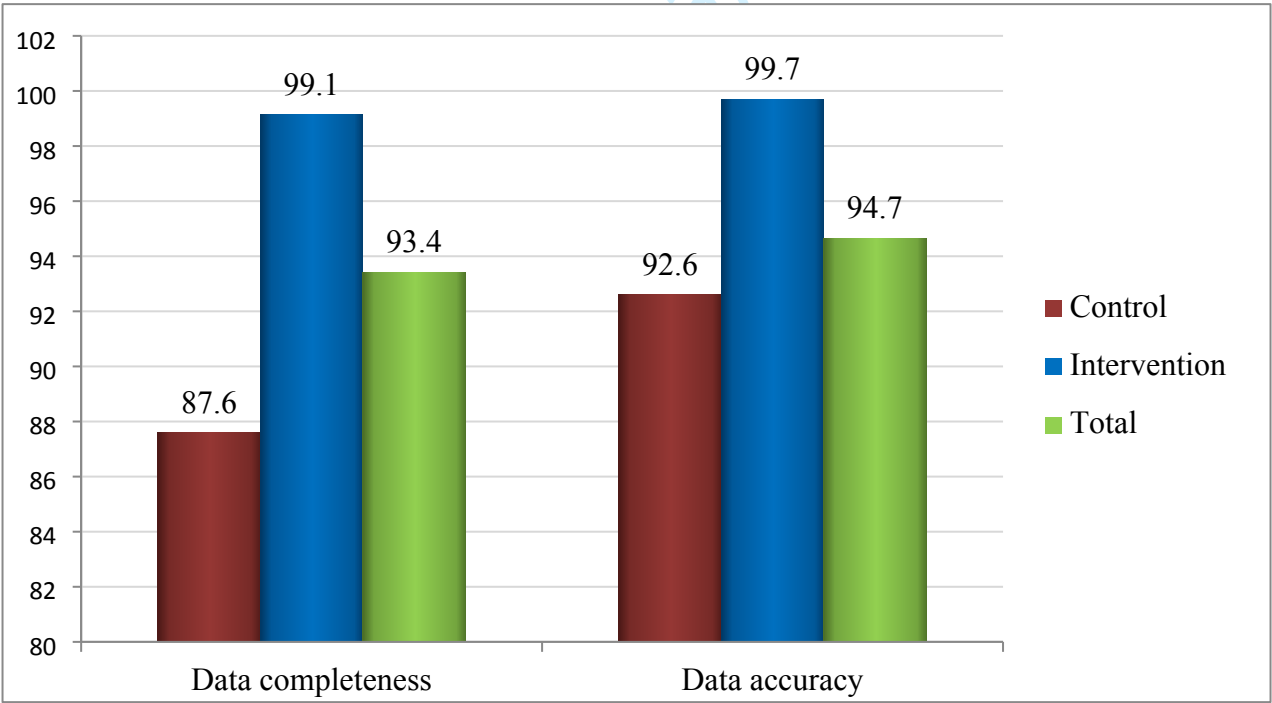


Figure 3. Data completeness and accuracy among treatment groups in health institutions of Gofa Zone, Southern Ethiopia, 2024 (n = 72).

Data quality practice score

At baseline, the mean score for data quality perception was 2.32 (95% CI: 2.25, 2.35; SE = 0.02), with a total of 59% respondents scoring at or above the average, and thus categorized as having a “good” perception of data quality practice. At end-line, a total of 77.3% of respondents had a good perception of data quality, with a mean data quality perception score of 3.13 (95% CI: 3.05, 3.21; $p < 0.001$). Overall, 68.7% of respondents had a good perception of data quality, having a mean score of 2.57 (95% CI: 2.52, 2.62; $p = 0.001$), with intervention institutions scoring significantly higher (mean = 2.77; 95% CI: 2.70, 2.83; SE = 0.02) than control institutions (mean = 2.38; 95% CI: 2.31, 2.44; SE = 0.02) (Figure 4).

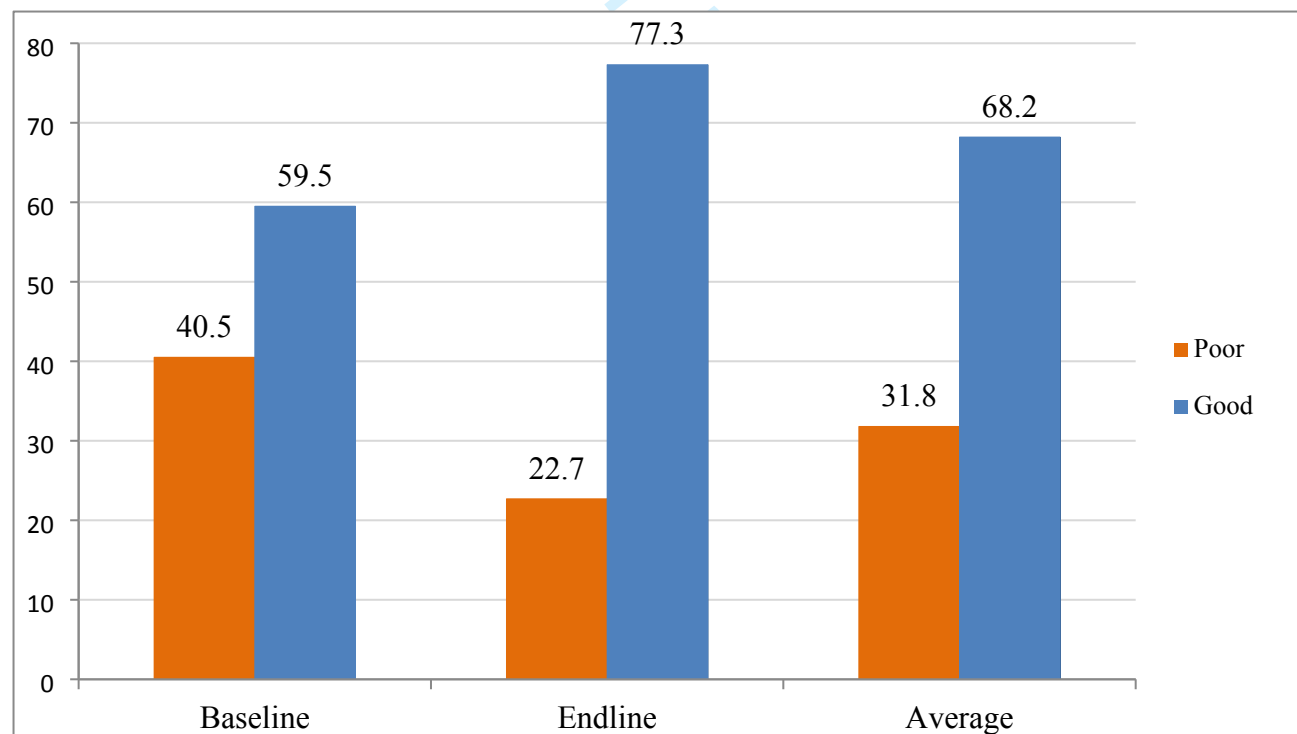


Figure 4. The average change in data quality score from baseline to end-line among health workers in public health institutions of Gofa zone, Southern Ethiopia, (n= 304).

Data quality and associated factors

In the bi-variable analysis, treatment group, time, encouraging system for good performance, ease of data management, level of information use, availability of written guideline on RHIS, receive supportive supervision on RHIS, training on RHIS, the combined effects of encouraging system for good performance and Training on RHIS, availability PMT system, receive feedback on RHIS, consistently use standard tools, and availability of rewarding system were significantly predicted the change in the data quality practice. However, in multivariable analysis predictors like treatment group, time, encouraging system for good performance, ease of data management, level of information use, availability of written guideline on RHIS, consistently use standard tools and the combined effects of encouraging system for good performance and training on RHIS, were significantly predicted the change in the data quality practice among the treatment groups (**Table 2**)

The overall effect of intervention. On average, health workers in the treatment group scored 0.17 points higher in data quality practice than those in the control group, after adjusting for other predictors (95% CI: 0.05, 0.30; p = 0.007). Similarly, a unit increases in time increases data quality score by 0.29 units (95% CI: 0.17, 0.41, p < 0.001).

Training and encouragement practice: A unit increase in the encouragement system for good performance in RHIS significantly increases the data quality score by 0.53 units (95% CI: 0.29, 0.76; p < 0.001). The combined intervention of training on RHIS and an encouragement system for good performance improved data quality by 0.44 units (95% CI: 0.23, 0.65; p = 0.010).

Ease of data management skill: it is a technical factor that evaluated the individual perception of how they feel about simplicity of operating data management process. In this regard, having data management skill significantly improves the data quality by 0.14 (95% CI=0.07, 0.22, $p < 0.001$).

Perceived level of information use: On average, when health workers have better perception on information utilization, their data quality practice significantly improve by 0.15 unit (95% CI: 0.08, 0.23), $p < 0.001$).

Availability of written guideline on RHIS: when there were written guidelines on RHIS in the department of the health workers, their perceived level of data quality practice increases ($\beta = 0.14$; 95% CI: 0.04, 0.24, $p = 0.005$).

Consistently using standard tools: a unit increase in consistence in use for standard tools increase data quality practice by 0.11 units (95% CI: 0.02, 0.21; $p = 0.023$).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

1 Table 2. General Linear Mixed Effect model estimation of data quality practice score with predictors among health workers in health
2 institutions of Gofa zone, South Ethiopia region, 2024 (n= 304).

Variables	Bivariate Model			Multivariate Model		
	β (95% CI)	SE	P	β(95% CI)	SE	p
Constant	2.69 (2.59, 2.79)	0.05	<0.001	1.50(1.27, 1.73)	0.12	<0.001
Treatment group	0.71 (0.61, 0.81)	0.05	<0.001	0.171 (0.05, 0.29)	0.06	0.007
Time	0.82 (0.70, 0.94)	0.06	<0.001	0.293 (0.17, 0.41)	0.06	<0.001
Encouraging system for good performance	1.35 (1.25, 1.45)	0.05	<0.001	0.530 (0.29, 0 .77)	0.12	<0.001
Ease of data management	0.39 (0.29, 0.47)	0.05	<0.001	0.142 (0.07, 0.22)	0.04	<0.001
Level of information use	0.41 (0.32, 0.50)	0.05	<0.001	0.152 (0.08, 0.23)	0.04	<0.001
Availability of written guideline on RHIS	0.47 (0.35, 0.58)	0.06	<0.001	0.144 (0.04, 0.24)	0.05	0.005
Receive supportive supervision on RHIS	0.51 (0.40, 0.62)	0.06	<0.001	-0.045 (-0.14, 0.09)	0.07	0.520
Training on RHIS	0.40 (0.30, 0.50)	0.05	<0.001	-0.042 (-0.15, 0.07)	0.06	0.460
Encouraging system for good performance x Training on RHIS	0.46 (0.23, 0.68)	0.11	<0.001	0.437 (0.23, 0 .65)	0.11	<0.001
Availability PMT system	0.17 (0.02, 0.31)	0.07	0.026	-0.072 (-0.19, 0.05)	0.06	0.229
Receive feedback on RHIS	0.28 (0.16, 0.39)	0.06	<0.001	-0.024 (-0.14, 0.09)	0.06	0.684
Consistently use standard tools	0.31 (0.19, 0.42)	0.06	<0.001	0.112 (0.02, 0 .21)	0.05	0.023
Availability of rewarding system	0.94 (0.82, 1.05)	0.06	<0.001	-0.004 (-0.16, 0.15)	0.08	0.962

Model fitness

Comparing Akaike's Information Criterion (AIC) and Bayesian Criterion (BIC) between the null and final models provides an indication of model fit. Lower values of the information criteria from the null to the final model suggest better-fitting models. On the other hand, the ICC measures the proportion of the total variance explained by the grouping factor. A higher ICC value ($> 5\%$) indicates that the grouping factor has a significant impact on the outcome variable (**Table 3**).

Table 3. Indicating the model fitness in the mixed linear effect model analysis among health workers in public health institutions of the Gofa zone, Southern Ethiopia, 2024 (n = 304).

Information Criteria	Null model	Final model
ICC (%)	14.55	8.20
AIC	1397.28	824.62
BIC	1414.84	903.61

Akaike's Information Criterion (AIC), Bayesian Criterion (BIC), Intraclass Correlation Coefficient (ICC)

Discussion

The aim of this study was to evaluate the effect of capacity development interventions on the performance of data quality in the RHIS among public health institutions. In a simple descriptive analysis, the components of data quality dimensions, report timeliness and data completeness improved over time from baseline to end-line within the intervention groups. The overall data quality score changed from baseline to end-line and demonstrated a significant difference among the groups. The intervention groups, time, encouragement for good performance, availability of written guidelines on RHIS, combined effects of training and supportive supervision or the encouragement system, as well as perceptions of the ease of data management and information

utilization practices, were significant predictors and changed from baseline to end-line or between the intervention and control groups.

Regarding data quality dimensions, report timeliness improved from baseline to end-line among the treatment groups, even though it did not meet the national MOH standard of 90%. This result aligns with a study conducted at Metema Primary Hospital in Northwest Ethiopia, where report timeliness increased following the implementation of intervention packages such as training, supportive supervision, and feedback (19) and Philippines (30). The interventions we applied; such as training, supervision, mentorship, motivation, and monitoring and evaluation on RHIS, might have helped build the capacity and knowledge of health workers, motivate staff, and enhance data flow and submission processes.

Similarly, data completeness also showed significant improvement following the implementation of the interventions. A similar finding was reported in the Oromia region of Ethiopia, where data completeness increased from about 42% before the intervention to 100% post-intervention. The intervention included here were discussions and action plan preparation with the management team at district level, on-the-job training, supportive supervision, data audits, and performance review meetings (2). Another study in Nigeria also reported improvement in changes in data completeness (10). A possible explanation for the association could be the strategies we implemented, such as standardizing data collection and reporting tools, training staff, validating data, conducting regular supportive supervision, offering timely feedback, and closely monitoring performance, all of which may have impacted overall performance. Together, these actions could also have strengthened the data management process, leading to more complete and reliable health data in source documents and reporting formats, which may have contributed to improvements in data completeness and overall data quality.

In this study, the overall ACIDI intervention recorded a significant difference among the treatment groups, with greater improvement observed in the treatment groups compared to the control groups. Similarly, the intervention was found to be effective from baseline to end-line. Studies in the Oromia region of Ethiopia (2), Amhara region of Northwest Ethiopia (19), low- and middle income countries (31) showed improvement in the performance following the implementation of a combination of interventions, rather than single interventions. These packages of interventions included preparing an action plan of activities, on-the-job training, supportive supervision, provision of feedback, data audits, and performance review meetings (32). The combination of interventions greatly impacts helping to address multiple dimensions of data quality (33). Training provides the essential knowledge and skills necessary for accurate data collection and reporting (14,20). Supportive supervision ensures consistent, on-the-ground guidance to address challenges and reinforce correct practices (7). Feedback highlights errors or weaknesses, offering opportunities to correct and improve data practices in real time (5). Review meetings serve as platforms for teams to assess performance, share lessons learned, and engage in collaborative problem-solving. Motivation creates a culture of accountability and encourages individuals to take ownership of the quality of the data they produce (34).

Encouragement for good performance in RHIS was among the predictors that significantly improved after application of the implementation in this study. Studies in Ethiopia and other parts of the globe have indicated that data quality practice is highly associated with institutional management support and encouragement (25,35). Our intervention has involved the heads of institutions and departments, whose encouragement and support could have made staff more effective and dedicated by addressing resource limitations and creating a supportive work

environment in which staff feel safe to try new approaches and perform better through boosted morale.

As indicated in this study, when training is combined with encouragement and support from managing bodies, there is a statistically significant improvement in the perceived quality of routine data. Some descriptive studies have revealed the association of training and data quality practice (20,36). A study in Ghana has revealed that a combination of interventions is necessary to achieve the intended outcome of data quality (37). Giving training to the healthcare workforce, especially to the heads of health institutions and leaders of departments, help them gain a deeper understanding of the importance of data quality in healthcare decision-making (38). This awareness encourages a culture of prioritizing accurate, reliable, and consistent data at all levels of the institution.

The perceived ease of data management skills and processes is a significant predictor of data quality practices that improved from before to after the intervention in the study. A study in in Oromia Special Zone, Amhara region of Ethiopia indicated a relationship between competency and data quality in RHIS (39). Another study in Massaguet district of Chad also indicated an association between the presence of a health technician and staff dedicated to data management and data quality in the HMIS (40). The perceived ease of data management skills and processes influences user engagement and motivation (41). When health workers and managers perceive data management processes as easy to understand and perform, they are more likely to engage consistently in proper data recording and reporting, feel confident in handling data, be motivated to ensure accuracy and completeness, and complete records on time while avoiding errors (42).

The perceived level of information utilization is one of the predictors of data quality that showed improvement from baseline to end-line among the groups. A PRISM framework analysis revealed that organizational and behavioral determinants, such as data use, are key factors influencing data

quality (43). The likely explanation for this association is that using data in decision-making fosters accountability and ownership among health workers by reinforcing their responsibility for producing quality data. Regular data reviews help identify and correct errors, thereby strengthening validation practices. This promotes a culture of continuous data quality improvement, motivates accurate reporting, and drives ongoing improvement through targeted actions based on identified gaps (44).

Using written guidelines in RHIS is an independent predictor of improved data quality practice in the study. As indicated by the Health Metrics Network, written guidelines promote standardization of data collection and reporting by ensuring uniform methods, definitions, and indicators, thereby minimizing errors and inconsistencies (3).

Consistency in using standard tools was a significant predictor of data quality practice in the study. A qualitative study conducted in Eastern Ethiopia revealed that the lack of standard forms was one of the barriers to data quality (34). According to a report in South Africa, the standardization of routine data collection and reporting tools strengthens the system, supports a nationwide common platform, and reduces fragmentation in the health information data management system (45). A possible explanation for the association of standardization tool utilization and data quality is that consistency in using documentation tools in RHIS is an important process that supports the regular, standardized, and systematic use of approved tools. These tools include registers, tally sheets, reporting forms, and electronic health records by health workers during service delivery, data recording, and reporting (46,47).

Strengths and weaknesses of the study

Strength of the study is the use of a cluster RCT, which helps minimize contamination by randomizing entire groups rather than individuals, reducing bias from cross-group interactions. Cluster RCTs reflect real-world implementation, are practical for group interventions, and improve statistical power. They are cost-effective, ethically sound, control for group-level confounders, and allow for long-term impact measurement. Combining general linear mixed models with ITT enhances statistical power, minimizes bias, handles missing data, improves generalizability, ensures ethical transparency, addresses confounding, and provides accurate, real-world effect estimates.

A limitation of the study is that there is a geographical proximity between some districts, with limited buffer zones, which may compromise the risk of contamination, even though there is no contact between institution to institution. During sample size calculations, the assumption of equal cluster size was considered, but in practice, the number of health workers selected varied among health institutions.

Conclusion

The ACDIs implemented in this study were found to be highly effective in influencing and bringing about the desired changes in data quality improvement. An encouraging system for good performance in RHIS, ease of data management, perceived level of information use, availability of written guidelines on RHIS, the synergetic effects of receiving supportive supervision on RHIS and training on RHIS, as well as the joint effects of an encouraging system for good performance and RHIS training, were significant predictors of data quality practices. Providing training for large groups of managers and health workers, and integrating RHIS training with supportive supervision are recommended strategies to improve data quality in health. Enhancing the data management skills of health workers and cultivating a culture of information use are also recommended

interventions. Therefore, these approaches can be scaled and utilized in similar settings to improve data quality practices.

List of abbreviations

ACDI = Augmented Capacity Development Interventions

ANC = Antenatal Care

CONSORT = Consolidated Standards of Reporting Trials

DHIS2 = District Health Information Software, version 2

DQA = Data Quality Assessment

HMIS = Health Management Information System

MCH = Maternal and Child Health,

OPD = Outpatient Department

PMT = Performance Monitoring Team

PRISM = Performance of Routine Information System Management

RHIS = Routine Health Information System

WHO = World Health Organization

Declarations

Acknowledgements

We are grateful to Hawassa University, College of Health and Medical Sciences, School of Public Health and former Southern National Nationalities and Peoples Regional Health Bureau for giving us the opportunity to conduct this study. We would also like to thank the Doris Duke Charitable Foundation (DDCF) for financially supporting the data collection process of this study.

Authors' contributions

BK involved in designing the study, data collection, data analysis, data interpretation, and manuscript write up. DK, AG, and KG were participated in the design of the study, critically reviewed and revised the manuscript. All authors have approved the final version of the manuscript.

Statements and declarations

Ethical considerations

Ethical approval of the protocol for this study was received from the institutional Review Board of the College of Medicine and Health Sciences, Hawassa University with the Reference No. of IRB/183/14 and date 08/06/2022. Approval letter was received from former Southern Nations, Nationalities and Peoples Region (SNNPR) Health Bureau. Permission letter was also obtained from the Gofa Zone Health Department, District Health Offices and each of respective health facilities. All procedures were conducted based on the voluntary participation of the study participants in compliance with the Helsinki Declaration of ethical principles. The study protocol was registered on 14 December 2022 at the Pan African Clinical Trial registry with ID number of PACTR202212472091194. The control institutions have received ACDI manuals after endpoint data collection.

Consent to participate

Written informed consent to participate in this study was obtained from all participants after providing full information about the purpose, procedures, potential risks, and benefits of the study. Participation was voluntary, and confidentiality was assured.

Consent for publication

Not applicable

Declaration of conflicting interests

The Authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding statement

Financial support for fieldwork of this study was provided **entirely** by a **grant from the Doris Duke Charitable Foundation (Grant Number: 2017187)**. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

Data availability

The data sets utilized in this study are available from the corresponding author and provided on reasonable request of authorized personnel.

ORCID iDs

Bedilu Kucho Doka, MPH  <https://orcid.org/0009-0003-0334-2305>

Dejene Hailu, PhD  <https://orcid.org/0000-0001-7870-1753>

Abebaw Gebeyehu, PhD  <https://orcid.org/0000-0003-2954-7379>

Keneni Gutema, PhD  <https://orcid.org/0000-0003-2697-7242>

References

1. Dagneu E, Woreta SA, Shiferaw AM. Routine health information utilization and associated factors among health care professionals working at public health institution in North Gondar, Northwest Ethiopia. BMC Health Services Research. 2018 Sep 4;18(1):685.
2. Dufera FN. Evaluation of HMIS Data Quality and Information Use Improvement for Local Action-Oriented Performance Monitoring in Beghi District in West Wollega, Oromia, Ethiopia. Journal of Health, Medicine and Nursing. 2018;50(0):47–57.
3. Health Metrics Network. Framework and standards for country health information systems. 2008.
4. Joshi C, Torvaldsen S, Hodgson R, Hayen A. Factors associated with the use and quality of antenatal care in Nepal: a population-based study using the demographic and health survey data. BMC Pregnancy Childbirth. 2014 Mar 3;14(1):94.

5. Chekol A, Ketemaw A, Endale A, Aschale A, Endalew B, Asemahagn MA. Data quality and associated factors of routine health information system among health centers of West Gojjam Zone, northwest Ethiopia, 2021. *Front Health Serv.* 2023 Mar 24;3:1059611.

6. Singh M, Goel MK, Behera BK, Punia JS, Labani S. Health Management Information System Data Quality under NRHM in District Sonipat, Haryana. *International Journal of Health Sciences.* 2016;(9).

7. Getachew N, Erkalo B, Garedew MG. Data quality and associated factors in the health management information system at health centers in Shashogo district, Hadiya zone, southern Ethiopia, 2021. *BMC Medical Informatics and Decision Making.* 2022 Jun 15;22(1):154.

8. Mucee EM, Odhiambo-Otieno GW, Kaburi W, Kinyamu RK. Routine Health Management Information Use in the Public Health Sector in Tharaka Nithi County, Kenya. 2016;2(3):13.

9. Tilahun B, Derseh L, Atinafu A, Mamuye A, Mariam TH, Mohammed M, et al. Level and contributing factors of health data quality and information use in two districts in Northwest Ethiopia: social-ecological perspective. *BMC Med Inform Decis Mak.* 2021 Dec;21(1):373.

10. Bhattacharya AA, Allen E, Umar N, Audu A, Felix H, Schellenberg J, et al. Improving the quality of routine maternal and newborn data captured in primary health facilities in Gombe State, Northeastern Nigeria: a before-- and--after study. *Open access.* :10.

11. Gebreslassie AA, Below MT, Ashebir MM, Gezae KE, Chekole MK, Gebreslassie AA, et al. Enhancing health facility-based data quality and use for decision making at primary health care units to improve health service delivery of maternal newborn child and adolescent health, Tigray Ethiopia 2018. *Archives of Community Medicine and Public Health.* 2020 Apr 6;6(1):031–5.

12. Traoré A, Kouyaté M, Maiga A, Ouatarra A, Dembélé I, Doumbia A, et al. Assessment of the Performance of Routine Health Information System Management in Mali (2018). :1.

13. Kebede M, Adeba E, Chego M. Evaluation of quality and use of health management information system in primary health care units of east Wollega zone, Oromia regional state, Ethiopia: *BMC Med Inform Decis Mak.* 2020 Dec;20(1):107.

14. Teklegiorgis K, Tadesse K, Mirutse G, Terefe W. Level of data quality from Health Management Information Systems in a resources limited setting and its associated factors, eastern Ethiopia. *SA Journal of Information Management.* 2016 Aug 10;18(1):8.

15. Endriyas M, Alano A, Mekonnen E, Ayele S, Kelaye T, Shiferaw M, et al. Understanding performance data: health management information system data accuracy in Southern Nations Nationalities and People’s Region, Ethiopia. *BMC Health Services Research.* 2019;

16. Rendell N, Lokuge K, Rosewell A, Field E. Factors That Influence Data Use to Improve Health Service Delivery in Low- and Middle-Income Countries. *Global Health: Science and Practice.* 2020 Oct 1;8(3):566–81.

17. Samuel K Cheburet1,* and G W. Odhiambo-Otieno. Performance of Routine Health Information System Management in Liberia. :41.

18. Aqil A, Ávila JEH, Mejía LSP, Parbul AS, Plaza B, Wilson N, et al. Guanajuato SINAIS Assessment. :44.
19. Gonete TZ, Yazachew L, Endehabtu BF. Improving data quality and information utilization at Metema Primary Hospital, Amhara national regional state, Northwest Ethiopia 2018: Capstone project. *Health Informatics J*. 2021 Jul;27(3):146045822110431.
20. Shama AT, Roba HS, Abaerei AA, Gebremeskel TG, Baraki N. Assessment of quality of routine health information system data and associated factors among departments in public health facilities of Harari region, Ethiopia. *BMC Med Inform Decis Mak*. 2021 Dec;21(1):287.
21. Chanyalew MA, Yitayal M, Atnafu A, Tilahun B. Routine health information system utilization for evidence-based decision making in Amhara national regional state, northwest Ethiopia: a multi-level analysis. *BMC Med Inform Decis Mak*. 2021 Dec;21(1):28.
22. World Health Organization. GUIDE TO THE HEALTH FACILITY DATA QUALITY REPORT CARD [Internet]. WHO; Available from: https://www.who.int/healthinfo/DQRC_Indicators.pdf
23. Many A, Nielsen P. Reporting Practices and Data Quality in Health Information Systems in Developing Countries: An Exploratory Case Study in Kenya. :13.
24. Njuguna C, Vandi M, Mugagga M, Kanu J, Liyosi E, Chimbaru A, et al. Institutionalized data quality assessments: a critical pathway to improving the accuracy of integrated disease surveillance data in Sierra Leone. *BMC Health Serv Res*. 2020 Dec;20(1):724.
25. Kebede M, Adeba E, Chego M. Evaluation of quality and use of health management information system in primary health care units of east Wollega zone, Oromia regional state, Ethiopia: *BMC Medical Informatics and Decision Making*. 2020 Jun 12;20(1):107.
26. Organization WH. Data quality review: module 1: framework and metrics [Internet]. World Health Organization; 2017 [cited 2020 Nov 15]. Available from: <https://apps.who.int/iris/handle/10665/259224>
27. Organization WH. Data quality review: module 3: data verification and system assessment [Internet]. World Health Organization; 2017 [cited 2020 Nov 15]. Available from: <https://apps.who.int/iris/handle/10665/259226>
28. MEASURE Evaluation. Performance of Routine Information System Management (PRISM) Toolkit: PRISM Tools [Internet]. Chapel Hill, NC, USA: MEASURE Evaluation.; 2019. Available from: <https://www.measureevaluation.org/prism>
29. Fretheim A. Interrupted time-series analysis yielded an effect estimate concordant with the cluster-randomized controlled trial result. *Journal of Clinical Epidemiology*. 2013;5.
30. Murai S, Ventura RJC, Gaite JT. Timeliness of reporting process in the national routine health information system: The case of 19-year experience of Field Health Services Information System in Palawan, the Philippines. Grépin KA, editor. *PLoS ONE*. 2022 Feb 25;17(2):e0264681.

1
2
3 31. Lemma S, Janson A, Persson LÅ, Wickremasinghe D, Källestål C. Improving quality and use of routine
4 health information system data in low- and middle-income countries: A scoping review. Francis JM,
5 editor. PLoS ONE. 2020 Oct 8;15(10):e0239683.
6
7 32. Wagenaar BH, Gimbel S, Hoek R, Pfeiffer J, Michel C, Manuel JL, et al. Effects of a health information
8 system data quality intervention on concordance in Mozambique: time-series analyses from 2009–
9 2012. Popul Health Metrics. 2015 Dec;13(1):9.
10
11 33. Lestari D, Tamtomo D, Murti B. Meta Analysis: Effects of Supervision Support, Feedback, and Training
12 on Utilization of Routine Health Information System in Health Workers. Journal of Health Policy and
13 Management. 2023 Sep 16;8(3):201–13.
14
15 34. Tolera A, Firdisa D, Roba HS, Motuma A, Kitesa M, Abaerei AA. Barriers to healthcare data quality and
16 recommendations in public health facilities in Dire Dawa city administration, eastern Ethiopia: a
17 qualitative study. Front Digit Health. 2024 Mar 14;6:1261031.
18
19 35. Gobena T, Berhanie D, Shore H, Kenay A, Wondirad Y, Ayanle M. Barriers and Enhancers of Data
20 Quality in Health Sector of Somali Regional State, Eastern Ethiopia. Ethiop J Health Dev.
21
22 36. Morris LD, Nyongesa MW, Sokiri TD. Factors influencing data quality in routine health information
23 systems in Maridi county, South Sudan. S Afr j inf manag [Internet]. 2024 Aug 21 [cited 2025 Apr
24 9];26(1). Available from: <https://sajim.co.za/index.php/sajim/article/view/1856>
25
26 37. Evaluating the Health Information system implementation and utilization in healthcare delivery.
27 Health Informatics Journal.
28
29 38. Tilahun B, Hailemariam T, Endehabtu BF, Atnafu A, Muluneh D, Mengesha M, et al. Capacitating
30 Health Sector Leaders to Improve Healthcare Data Quality and Use in Assosa District: Implementation
31 Research. Ethiop J Health Dev.
32
33 39. Molla A, Hayelom M, Adamu K, Mihiretu MM, Adem YF. Health data management practice and
34 associated factors among health professionals working in public health facilities in Oromia Special
35 Zone, Amhara, Ethiopia: a cross-sectional study. bmjph [Internet]. 2024 Jul 31 [cited 2025 Apr
36 10];2(1). Available from: <https://bmjpublichealth.bmj.com/content/2/1/e000807>
37
38 40. Moukénét A, de Cola MA, Ward C, Beakgoubé H, Baker K, Donovan L, et al. Health management
39 information system (HMIS) data quality and associated factors in Massaguet district, Chad. BMC
40 Medical Informatics and Decision Making. 2021 Nov 22;21(1):326.
41
42 41. Abebe H, Legesse M, Godie Y, Birhanu D, Belege F, Menbere F, et al. Health Management Information
43 System Data Quality and its associated factors in Addis Ababa Public Hospitals, Ethiopia, 2022.A cross-
44 sectional study [Internet]. medRxiv; 2025 [cited 2025 Apr 12]. p. 2025.01.19.25320816. Available
45 from: <https://www.medrxiv.org/content/10.1101/2025.01.19.25320816v1>
46
47 42. Muhoza P, Saleem H, Faye A, Tine R, Diaw A, Kante AM, et al. Behavioral Determinants of Routine
48 Health Information System Data Use in Senegal: A Qualitative Inquiry Based on the Integrated
49 Behavioral Model. Global Health: Science and Practice [Internet]. 2022 Jun 29 [cited 2025 Apr
50 10];10(3). Available from: <https://www.ghspjournal.org/content/10/3/e2100686>
51
52
53
54
55
56

- 1
2
3 43. Aqil A, Lippeveld T, Hozumi D. PRISM framework: a paradigm shift for designing, strengthening and
4 evaluating routine health information systems. *Health Policy and Planning*. 2009 May 1;24(3):217–28.
5
6 44. Nutley T, Reynolds HW. Improving the use of health data for health system strengthening. *Glob*
7 *Health Action* [Internet]. 2013 Feb 13 [cited 2020 Oct 26];6. Available from:
8 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3573178/>
9
10 45. Jacucci E, Shaw V, Braa J. Standardization of health information systems in South Africa: The challenge
11 of local sustainability. *Information Technology for Development*. 2006 Jul;12(3):225–39.
12
13 46. Haftu B, Taye G, Ayele W, Habtamu T, Biruk E. A mixed-methods assessment of Routine Health
14 Information System (RHIS) Data Quality and Factors Affecting it, Addis Ababa City Administration,
15 Ethiopia, 2020. *Ethiop J Health Dev*. 2021 Jul;10.
16
17 47. Rumisha SF, Lyimo EP, Mremi IR, Tungu PK, Mwingira VS, Mbata D, et al. Data quality of the routine
18 health management information system at the primary healthcare facility and district levels in
19 Tanzania. *BMC Med Inform Decis Mak*. 2020 Dec;20(1):340.
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

Supplementary File: Consolidated Standards of Reporting Trials (CONSORT) for Extension to Cluster Randomized Trials Completed Checklist

Manuscript Submission ID: f09209cb-0d88-45e6-bfce-7108ef2d884f

Manuscript Title: Effect of Augmented Capacity Development Interventions (ACDI) on the performance of data quality in the Routine Health Information System (RHIS) among health workers in public health institutions of Gofa Zone, Southern Ethiopia: a cluster randomized controlled trial

Table | CONSORT 2010 completed checklist of information to include when reporting a cluster randomized trial

Section/topic and item No	Standard checklist item with Extension for cluster designs	Authors' description	Page No
Title and abstract			
1a	Identification as a randomized trial in the title (Identification as a cluster randomized trial)	The cluster randomized trial design was indicated in the title to ensure appropriate indexing of the study in databases. It was also included as a keyword in the manuscript.	1
1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	As per the CONSORT guidelines, we have used a structured summary in the abstract, including the methods with trial design, results, and conclusions. The CONSORT extension for abstracts	2_3

		of cluster randomized trials was also taken into consideration.	
Introduction			
Background and objectives:			
2a	Scientific background and explanation of rationale for using a cluster design	The scientific background of the study was concisely presented. We provided a brief explanation for using the cluster design, highlighting the limitations of previous studies that failed to account for variations in outcomes among health institutions.	4_6
2b	Specific objectives or hypotheses	The specific aim of this study was presented in the final paragraph of the 'Introduction section.'	6
Methods			
Trial design:			
3a	Description of trial design (such as parallel, factorial) including allocation ratio (Definition of cluster and description of how the design features apply to the clusters)	A two-arm, parallel-group, cluster-randomized controlled trial was implemented in the study, as described in the trial design subsection of the Methods. However, the allocation ratio of clusters and individual participants is stated in the 'sampling procedures' and 'randomization' subsection of the Methods section within the manuscript.	6
3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	No significant changes were made to the methods after the implementation of the intervention began.	N/A
Participants:			
4a	Eligibility criteria for participants [and/or clusters]	The eligibility of participants and clusters, along with the inclusion	6_7

		and exclusion criteria, was briefly stated in a dedicated 'Eligibility' subsection.	
4b	Settings and locations where the data were collected	The setting of the study was briefly discussed in the first part of the Methods section.	6
Interventions:			
5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered [Whether interventions pertain to the cluster level, the individual participant level, or both]	The type, frequency, approaches, and dose of the intervention packages implemented in these studies, including training, supervision, mentorship, monitoring and evaluation, and motivation, were discussed in detail. Both health institutions, such as district health offices, hospitals, health centers, and health posts (clusters), as well as individual health workers, were targeted in the provision of the intervention.	11_13
Outcomes:			
6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed [Whether outcome measures pertain to the cluster level, the individual participant level, or both]	Data quality practice is the primary outcome of this study. It is a composite construct measured by the level of agreement on 11 items using a Likert scale, where a score of 5 represents 'Strongly Agree,' 4 represents 'Agree,' 3 represents 'Neutral,' 2 represents 'Disagree,' and 1 represents 'Strongly Disagree.' In this case, respondents' perceptions of how their institution performs in terms of data quality are assessed and measured at the individual level. On the other hand, other dimensions of data quality, such as completeness, timeliness, and accuracy, are considered secondary outcomes and measured at	9_11

		the cluster level.	
6b	Any changes to trial outcomes after the trial commenced, with reasons	No change in the outcome.	N/A
Sample size:			
7a	How sample size was determined [Method of calculation, number of clusters(s) (and whether equal or unequal cluster sizes are assumed), cluster size, a coefficient of intracluster correlation (ICC or k), and an indication of its uncertainty]	The study applied the assumptions of double population formula with the confidence level of 95%, marginal error of 5%, and intervention to control ratio of 1:1 to determine the sample size. The sample size was calculated by considering the percent of data quality in comparison group of 33% (20). Power of 90% was assumed to detect 30% difference in rates between the two groups. Since the study was a cluster design, ICC of 0.35 and average cluster size of 4.3 were utilized from previous related study (21). The design effect of 2.2 and non-response rate of 10% was considered. During sample size computation, the assumption of equal cluster size was applied, even though the actual size of the clusters varied among health institutions, which may be considered a limitation of the study.	7
7b	When applicable, explanation of any interim analyses and stopping guidelines	No interim analysis was conducted, as only baseline and end-line data were used. However, the data collected at baseline were used to identify system gaps, as the intervention included components based on gap-driven actions. The intervention was initially planned for	N/A

		eight months and was completed as scheduled, without any decision to stop implementation based on interim progress results.	
Randomization			
Sequence generation:			
8a	Method used to generate the random allocation sequence	Before the implementation of the randomization process, districts were first stratified by location type (i.e., urban or rural). Then, to reduce the risk of experimental contamination, districts were allocated using the block randomization procedure. Three (one urban and two rural) adjacent and contiguous districts were grouped into one block and the other three (one urban and two rural) districts were sorted in to the other block. Finally, the blocks were randomly selected and allocated into either intervention or control groups.	8_9
8b	Type of randomization; details of any restriction (such as blocking and block size) [Details of stratification or matching if used]	We applied stratified block randomization process by restricting adjacent districts to be categorized under similar groups.	
Allocation concealment mechanism:			
9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned [Specification that allocation was based on clusters rather than individuals and whether allocation concealment (if any) was at	In order to minimize the selection bias and ensure unpredictability, the assignment of the blocks to the either arm has been done by an independent researcher from Arbaminch University of Ethiopia, who was unaware of the study group assignments, applied sealed envelopes for the group allocation.	9

	the cluster level, the individual participant level, or both]		
Implementation:			
10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Discussed below	
10a	Who generated the random allocation sequence, who enrolled clusters, and who assigned clusters to interventions	An independent researcher from Arbaminch University of Ethiopia, who was unaware of the study group assignments, applied sealed envelopes for the group allocation. Then the research team enrolled participants, and assigned participants to interventions.	9
10b	Mechanism by which individual participants were included in clusters for the purposes of the trial (such as complete enumeration, random sampling)	All health workers serving in the intervention health institutions were considered eligible to receive the intervention. However, due to practical and financial constraints, we purposively selected heads of institutions and departments. In addition to these participants, health workers from key departments such as OPD and MCH were proportionally allocated and randomly selected.	8
10c	From whom consent was sought (representatives of the cluster, or individual cluster members, or both) and whether consent was sought before or after randomization	Ethical approval of the protocol for this study was received from the institutional Review Board of the College of Medicine and Health Sciences, Hawassa University with the Reference No. of IRB/183/14 and date 08/06/2022. Approval letter was	32

		received from former Southern Nations, Nationalities and Peoples Region (SNNPR) Health Bureau. Permission letter was also obtained from the Gofa Zone Health Department, District Health Offices and each of respective health facilities. Written informed consent was obtained from each study participant at both data collection points.	
Blinding:			
11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	In order to avoid any bias on study results, the outcome assessors were withheld about the interventions provided as they were deployed from unselected districts. The blinding of program implementers and study participants was not possible as they provide and receive the open-level trial. However, the control groups were kept unaware of what the intervention groups received.	9
11b	If relevant, description of the similarity of interventions	The intervention packages were designed and delivered similar across all clusters in terms of content, delivery approach, frequency, and duration, ensuring consistency in implementation and allowing for reliable comparison of outcomes across study groups.	11_13
Statistical methods:			
12a	Statistical methods used to compare groups for primary and secondary outcomes [How clustering was taken into account]	To compare groups for the primary and secondary outcomes, we used general linear mixed models (GLMM) with repeated measures. This approach allowed us to account for both the clustering of	14

		participants within health facilities and the correlation between repeated observations over time. Random effects for clusters were included to adjust for intra-cluster correlation, ensuring accurate estimation of intervention effects.	
12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	Adjusted analyses were conducted using general linear mixed models, controlling for potential confounding variables. These adjustments helped improve the precision of effect estimates and account for underlying differences between groups. Adjusted analyses were also applied in computing baseline characteristics.	16_18 & 24
Results			
Participant flow (a diagram is strongly recommended):			
13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome [For each group, the numbers of clusters that were randomly assigned, received intended treatment, and were analysed for the primary outcome]	The flow chart of study participants is presented in Figure 1 of the manuscript. The number of individuals and clusters randomly assigned and analyzed in each arm is clearly displayed in the diagram.	14_16
13b	For each group, losses and exclusions after randomization, together with reasons [For each group, losses and exclusions for both clusters and individual cluster members]	The exclusion criteria and loss to follow-up were clearly discussed in the manuscript, as indicated below: Newly established (2 health posts), nonfunctional (4 health posts) and privately owned health facilities were not considered in this	14_16

		study. The health workers who were not available during baseline data collection (5 health workers); who intended to leave the institution within eight months immediately prior to the baseline data collection (6 health workers); and who did not receive the intervention or dropped out at some point (13 health workers) were also excluded.	
Recruitment:			
14a	Dates defining the periods of recruitment and follow-up	<p>The dates of baseline and end-line data collection, as well as the intervention period, are defined in the 'Study Period' section of Methods in the manuscript.</p> <p>The baseline data were collected from April 1 to 30, 2023. The eight-month intervention was implemented from July 1, 2023 to February 29, 2024. The end-line data were collected from April 1 to 30, 2024.</p>	6
14b	Why the trial ended or was stopped	The trial ended after the initial eight-month implementation plan was completed.	N/A
Baseline data:			
15	A table showing baseline demographic and clinical characteristics for each group [Baseline characteristics for the individual and cluster levels as applicable for each group]	The baseline characteristics of the study are presented in Table 1 of the manuscript. The frequency of participants in each group is reported, along with p-values, adjusted for clusters.	16_18

Numbers analyzed			
16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups [For each group, number of clusters included in each analysis]	A total of 70 health institutions (6 districts, 2 hospitals, 18 health centers, and 44 health posts) as well as 291 health workers were included in the final data collection. However, we used intention-to-treat analysis, where all randomized clusters (72) and individuals (304) at the beginning of the study were considered in the final analysis.	16_18
Outcomes and estimation:			
17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval) [Results at the individual or cluster level as applicable and a coefficient of intracluster correlation (ICC or k) for each primary outcome]	The outcome variable is continuous, and we applied general linear mixed models for the repeated data. In this regard, the standardized coefficient along with its 95% confidence interval is presented. The intracluster correlation (ICC) of the null and final models is presented for the evaluation of model fit.	24
17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	N/A	
Ancillary analyses			
18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing prespecified from exploratory	Adjusted analyses were conducted using general linear mixed models, controlling for potential confounding variables. These adjustments helped improve the precision of effect estimates and account for underlying differences between groups. Adjusted analyses were also applied in computing baseline characteristics.	16_18 & 24

Harms:			
19	All important harms or unintended effects in each group	No harm or unintended results occurred.	N/A
Discussion			
Limitation			
20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	<p>The limitations of the study are presented in the Discussion section of the manuscript under the heading "Strengths and Weaknesses of the Study," as stated below.</p> <p>A limitation of the study is that there is a geographical proximity between some districts, with limited buffer zones, which may compromise the risk of contamination, even though there is no contact between institution to institution. During sample size calculations, the assumption of equal cluster size was considered, but in practice, the number of health workers selected varied among health institutions.</p>	29_30
Generalizability:			
21	Generalizability (external validity, applicability) of the trial findings [Generalizability to clusters and/or individual participants (as relevant)]	We believe that the finding is generalizable because we properly designed the study sample, utilized a gold-standard study design, implemented proper outcome measurement, and accurately applied the intervention delivery. Therefore, the study offers valuable evidence that reflects real-world conditions.	N/A
Interpretation:			

22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	Interpretation of this study is consistent with the results, involving drawing conclusions directly from the study's data, avoiding overgeneralizations or unsupported extrapolations. We balance and report the positive outcomes of intervention improvements against potential negative consequences with no associations. Additionally, we consider other relevant evidence or findings from similar studies to help contextualize the results within the broader research landscape. This comparison enhances confidence in the conclusions, ensuring the results are well-supported and not based on isolated or anomalous data, providing a comprehensive and accurate understanding.	25_29
Other information			
Registration:			
23	Registration number and name of trial registry	The study protocol of this study was registered at the Pan African Clinical Trial registry with ID number of PACTR202212472091194.	2
Protocol:			
24	Where the full trial protocol can be accessed, if available	The protocol of this study is available and can be accessed from the authors.	N/A
Funding			
25	Sources of funding and other support (such as supply of drugs), role of funders	The fieldwork of this study was supported by the Doris Duke Charitable Foundation through the Hawassa University Project. The	32

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

		funder has no role in design of the study, data collection process, data analysis, and publication of the article.	
--	--	---	--

For Peer Review