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 of cluster randomized trials was also taken into consideration. | 2\_3 |
| **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 and exclusion criteria, was briefly stated in a dedicated 'Eligibility' subsection. | | 6\_7 |
| 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 the cluster level. | | 9\_11 |
| 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. **S**ince 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 eight months and was completed as scheduled, without any decision to stop implementation based on interim progress results. | | N/A |
| **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 the cluster level, the individual participant level, or both] | | 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 |
| 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 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. | | | 32 |
| 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 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. | | 14 |
| 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 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. | | 14\_16 |
| Recruitment: | | | | | |
| 14a | Dates defining the periods of recruitment and follow-up | | 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.  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. | | 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 | | 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.  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. | | 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 funder has no role in design of the study, data collection process, data analysis, and publication of the article. | | 32 |