


Interventions to improve district-level routine health data in low-income and middle-income countries: a systematic review

Jieun Lee,^{1,2} Caroline A Lynch,³ Lauren Oliveira Hashiguchi,¹ Robert W Snow,^{4,5} Naomi D Herz,⁶ Jayne Webster,¹ Justin Parkhurst ,⁷ Ngozi A Erondy ^{1,8}

To cite: Lee J, Lynch CA, Hashiguchi LO, *et al*. Interventions to improve district-level routine health data in low-income and middle-income countries: a systematic review. *BMJ Global Health* 2021;**6**:e004223. doi:10.1136/bmjgh-2020-004223

Handling editor Edwine Barasa

► Additional online supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjgh-2020-004223>).

Received 19 October 2020
Accepted 20 May 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Ngozi A Erondy;
ngozierondy@gmail.com

ABSTRACT

Background Routine health information system(s) (RHIS) facilitate the collection of health data at all levels of the health system allowing estimates of disease prevalence, treatment and preventive intervention coverage, and risk factors to guide disease control strategies. This core health system pillar remains underdeveloped in many low-income and middle-income countries. Efforts to improve RHIS data coverage, quality and timeliness were launched over 10 years ago.

Methods A systematic review was performed across 12 databases and literature search engines for both peer-reviewed articles and grey literature reports on RHIS interventions. Studies were analysed in three stages: (1) categorisation of RHIS intervention components and processes; (2) comparison of intervention component effectiveness and (3) whether the post-intervention outcome improved above the WHO integrated disease surveillance response framework data quality standard of 80% or above.

Results 5294 references were screened, resulting in 56 studies. Three key performance determinants—technical, organisational and behavioural—were proposed as critical to RHIS strengthening. Seventy-seven per cent [77%] of studies identified addressed all three determinants. The most frequently implemented intervention components were ‘providing training’ and ‘using an electronic health management information systems’. Ninety-three per cent [93%] of pre–post or controlled trial studies showed improvements in one or more data quality outputs, but after applying a standard threshold of >80% post-intervention, this number reduced to 68%. There was an observed benefit of multi-component interventions that either conducted data quality training or that addressed improvement across multiple processes and determinants of RHIS.

Conclusion Holistic data quality interventions that address multiple determinants should be continuously practised for strengthening RHIS. Studies with clearly defined and pragmatic outcomes are required for future RHIS improvement interventions. These should be accompanied by qualitative studies and cost analyses to understand which investments are needed to sustain high-quality RHIS in low-income and middle-income countries.

Key questions

What is already known?

- Routine health information systems (RHIS) are foundational.
- The quality of data produced from RHIS in many low-income and middle-income countries remains inadequate.
- Since RHIS are composed of complex inputs and mechanisms, tackling data quality issues is also complex.

What are the new findings?

- There is remarkable diversity in the methods used to assess interventions aimed at improving RHIS at the district level and below.
- The most frequent types of studies included were quasi-experimental (n=22, 39%), case studies (n=11, 20%) and process evaluations (n=8, 14%).
- Seventeen discrete intervention components (eg, training, electronic health management information system) were identified.

What do the new findings imply?

- Holistic data quality interventions addressing technical, behavioural and organisational determinants of RHIS.
- More aligned methods for measuring, evaluating and benchmarking data quality are needed to better identify what interventions work to improve data quality.
- Continued attention and investment on RHI data quality, regardless of type of interventions used, is needed.

INTRODUCTION

Data regularly gathered by healthcare providers, referred to as routine health information (RHI), are used to inform countries on their health status, health service capacity and health resources needs. WHO identifies a health information system as a core health system building block.¹ An RHI system(s) (RHIS) facilitates ongoing facility-level

collection of health and management information and the regular aggregate reporting of this data between levels of the health system. A well-functioning RHIS can provide timely information on disease incidence, preventive intervention and treatment coverage, and risk factors to guide disease control strategies and epidemic responses. Furthermore, national disease burden estimates guide programmes and donor funding priorities in high disease burden low-income and-middle income countries (LMICs). While most countries have made progress in establishing well-functioning RHIS, investments in national health data systems in LMICs remains inadequate with still many gaps in data quality (DQ) and coverage.^{2,3} As a result, there is heavy reliance on mathematical modelling based on incomplete, and often inaccurate, data.⁴⁻⁷ In effect this means that decision-makers have only a partial view of their population's health situation, making it particularly challenging to accurately and effectively target health interventions that address their needs, particularly in LMICs^{8,9}

There have been successive calls for greater investments in RHIS over the last three decades⁹⁻¹² culminating in Chan *et al*¹³ leading a call to action in 2010 for a concerted and systematic effort by global partners to provide support to countries in strengthening their monitoring of progress and performance of DQ. Additionally, formal studies of how to improve routine health DQ, timeliness and fidelity have been guided by the Health Data Collaborative (HDC), following a high-level 2015 summit on Measurement and Accountability for Results in Health to 'improve efficiency and alignment

of technical and financial investments in health data systems through collective actions'.¹⁴

Previous reviews on interventions to improve RHIS focus on technical aspects of improving the quality of data, and much less on how the interventions address human-related factors and system processes (ie, data collection, data transmission) within the RHIS.^{8, 15, 16} Several authors have emphasised the need to focus on levels of the health system where data collection and entry occurs (ie, district levels and below). Moreover, the recent trend of decentralisation of government structures in the LMICs, means effective data use should take place at the district-equivalent level for timely and informed decision making.^{4,8,15}

The quality of data in routine health systems has received increased attention and focus over the last 10–15 years. However, we identified only one systematic review that described DQ interventions at peripheral levels of the health system.¹⁷ Here, we systematically review evidence to identify and compare the types and effectiveness of RHIS interventions at district and community levels in LMICs aimed at increasing DQ.

METHODS

Conceptual framework

We conceptualised DQ on the basis of the Performance of Routine Information Systems Management (PRISM) model presented by Aqil *et al*¹⁶ (figure 1). The PRISM framework describes DQ according to four dimensions: data accuracy; completeness; timeliness and relevance. We have also used the definitions suggested by Aqil *et al*

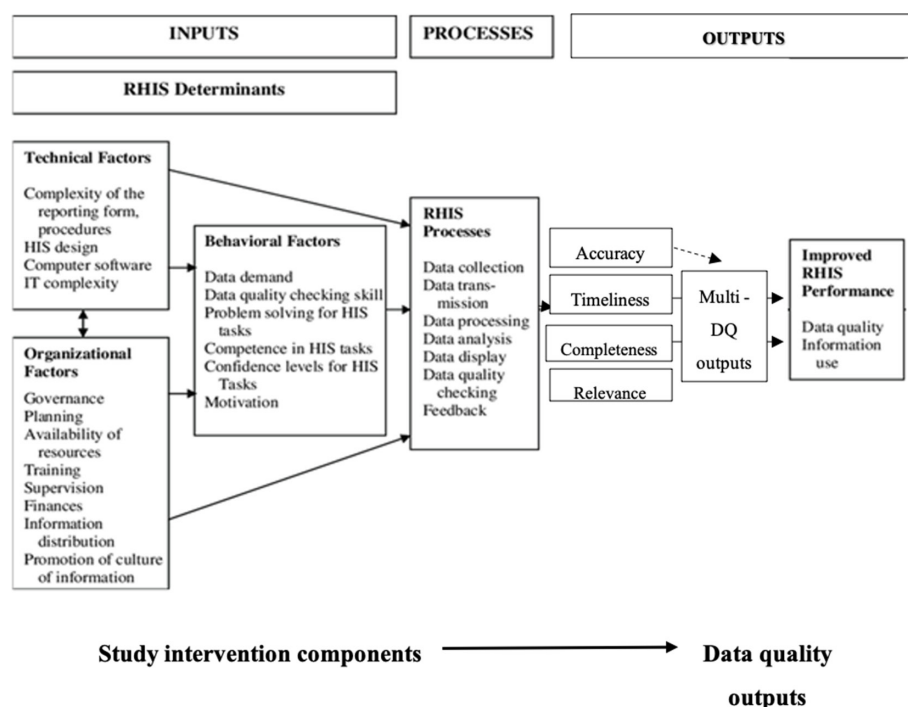


Figure 1 Performance of Routine Information Systems Management framework (adapted). DQ, data quality; HIS, health information system; RHIS, routine health information system.

(online supplemental file S1(1.1))¹⁶ to define each of the DQ outcomes. For this review, we refer to these dimensions collectively as ‘DQ outputs’.¹⁸

In addition to conceptualising DQ, the PRISM framework also identifies determinants of RHIS performance and processes for a functioning RHIS, with the aim of targeting interventions that could improve DQ outputs.¹⁸ The framework states that RHIS performance is affected by key processes (ie, data collection, data transmission, quality checking, etc), which are affected by technical, behavioural and organisational determinants (figure 1). Technical determinants concern the ‘know-how’ and ‘technology’, organisational determinants are the ‘rules, values and practices’ within the organisations and behavioural determinants are ‘users’ demand, confidence, motivation and competence’ that influences DQ and use.¹⁸ For this review, we examine interventions used to improve DQ and categorise those into organisational and behavioural determinants (organisational and behavioural determinants were combined to represent one category given the similarity and potential categorical overlap in most interventions) or technical components. The terms and concepts of the PRISM framework are further defined by Aqil *et al* in several articles and reports.^{16 18}

Search strategy

The systematic review was conducted using an adapted version of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.¹⁹ The literature search was conducted using five electronic databases for peer-reviewed articles, including the Medline In-Process, Embase, Global Health, Web of Science, Journal of Health Informatics for Developing Countries search engines. Other electronic sources including Popline and OpenGrey, MEASURE evaluation website, Routine Health Information Network resource library, Human Resources for Health digital library, John Snow Inc. website and Google search engine were searched for grey literature. Snow-balling technique was used to further identify relevant articles from the reference lists of selected studies. The first round of searches was conducted over 2 weeks and completed by 12 September 2017, and the review was then updated on 4 January 2021, covering Published and unpublished reports available between October and January 2021. (online supplemental file S1(1))

Search terms were composed of a combination of one or more of three major themes: RHIS and equivalent terms, LMICs as defined by the World Bank in 2017,²⁰ and DQ and terms equivalent (online supplemental file S1(1.2)). Results of the searches were exported to a reference management database. The literature search included articles published in English, French, Spanish and Portuguese. There was no specific limitation placed on the date of publication within databases used.

Study selection and outcomes

Selected articles included studies from LMICs that measured a change in DQ attributable to interventions being tested at levels of the health system referred to district level or below with a clearly stated aim to improve the DQ. Definitions of DQ used in each of the studies can be found in online supplementary file S1 (2.4). Trials, pre-post intervention, quantitative evaluations, case studies and qualitative studies were included if they involved a description of the intervention and the difference in (or perceived difference in), pre–post intervention DQ. We excluded animal health-related information systems as well as studies referring to improvements in data from surveys, censuses and vital registration generally. Articles without an intervention component specifically targeting DQ were excluded, as well as interventions that did not include district level and below interventions. Studies that did not include pre–post outcome measures, perceived changes in DQ, or some other comparison (such as a control site) were also excluded.

We selected the articles in three phases. Two reviewers (JL and LOH) completed the first round of article selection by reviewing titles. Second, a single reviewer (JL) screened all abstracts and second reviewer (LOH) screened a subset of the same abstracts to ensure agreement. At this stage, duplicates were identified and removed. For the third stage, one reviewer (JL) undertook a full-article review and again a subset was reviewed by a second reviewer (LOH) for quality control. Any discrepancies in coding of articles were resolved through discussion between reviewers. Where necessary opinion was sought from a third reviewer (NAE). Data extraction was carried out during the full-text screening stage. Major themes included characteristics of the selected studies (setting, study design, objectives, duration, study population), characteristics of the intervention (determinants and processes addressed, details of the intervention components) and the outcomes. During the updated search in January 2021, JL, CAL and NAE conducted the titles screening JL, CAL, NDH and NAE conducted abstract and full-text screening.

Data synthesis

For the quantitative studies, interventions were synthesised using three methods. First, we sought to provide an overview of included studies with a description of all DQ outputs measured, processes addressed and types of intervention components (figure 2). Second, we conducted a further analysis for quantitative studies to stratify the level of effectiveness and to highlight interventions that had a larger effect on DQ outcomes. For this, we first applied a more stringent standard of DQ with $\geq 80\%$ for accuracy, completeness and timeliness post-intervention. A threshold of $\geq 80\%$ was used, based on the only widely used standard available: the WHO performance standard for integrated disease surveillance and response systems.²¹ We refer to studies with improved DQ output and reached $\geq 80\%$ as the studies that are ‘above

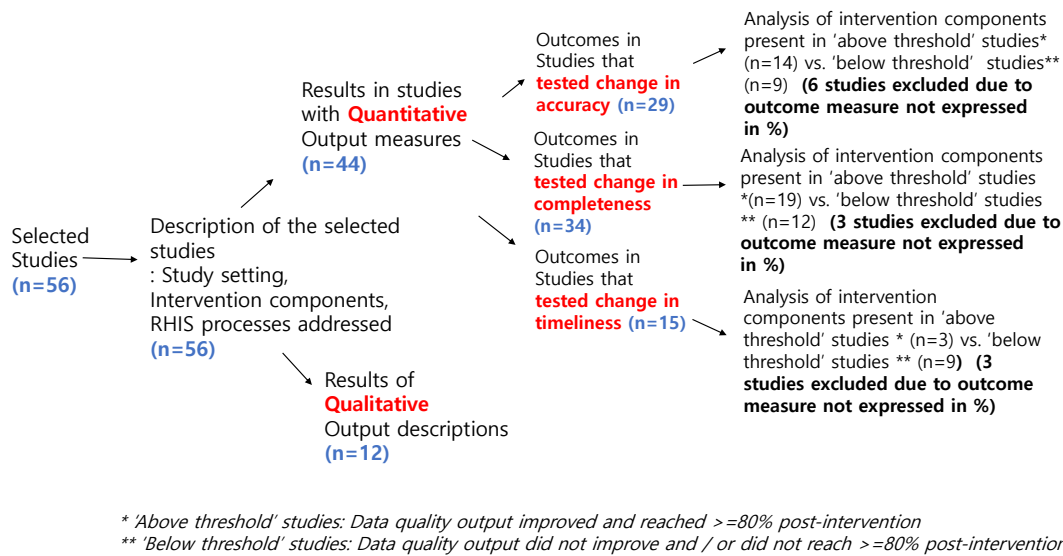


Figure 2 Flow chart showing outcomes of selected studies (quantitative and qualitative). * 'Above threshold' studies: Data quality output improved and reached $\geq 80\%$ post-intervention ** 'Below threshold' studies: Data quality output did not improve and / or did not reach $\geq 80\%$ post-intervention. RHIS, routine health information system.

threshold' for each DQ outputs (accuracy, completeness and timeliness), and the studies with unimproved DQ output and/or did not reach 80% as the studies that are 'below threshold'. For this analysis, we excluded studies for which there was no per cent output expressed. Third, we compared the frequency of different intervention components used and RHIS processes addressed in 'above threshold' and 'below threshold' studies for each of the DQ outputs, to identify what intervention components were more frequently used in studies with outcomes that improved and reached above threshold. In order to ascertain which intervention components (eg, training, task-shifting, meetings) may have contributed toward higher accuracy, we ranked each component by difference between below threshold and above threshold studies (ie, difference in proportion of number of studies in below threshold studies that measured accuracy and implemented training/total number of below threshold studies that measured accuracy vs number of studies in above threshold studies that measured accuracy and implemented training/total number of above threshold studies that measured accuracy). We also categorised studies that implemented more than four intervention components or addressed more than three processes as part of this comparison, to assess whether interventions with more components were more effective.

We assessed the qualitative studies separately—we focused on the captured determinants and process of these studies. We did not include the results of the qualitative studies in assessing whether there was improvement in accuracy, completeness, timeliness or when assessing whether the improved outcome reached the threshold or not.

For simplicity, we renamed intervention components that had the same name but differed in terms of their intensity. For example, some studies reported 1–5 days training courses, while others described training that took >12 weeks and included follow-up on-site training. We categorised the former as a 'training' intervention component and the latter as 'enhanced training' component. We used the same approach when differentiating between standard 'supervision' and 'enhanced supervision'. The latter included more regular (ie, daily, weekly, bimonthly and/or online technical support) while the former was less frequent. Where studies describe having meetings to discuss data, dissemination meetings or quarterly meetings, we aggregate those under the intervention component 'meetings'. DQ checks differ from DQAs (DQ Assessments), as DQAs refer to using specific tools (ie, Routine Data Quality Assessment) tool designed by MEASURE evaluation) to assess DQ in a routine manner, whereas DQ checks refer to data checking functions embedded within the electronic system or a stakeholder within the RHIS manually checking to capture and correct errors and incomplete fields. Lastly, database harmonisation refers to integrating multiple electronic health management information systems (eHMIS) to a single system.

Patient and public involvement

This systematic literature review did not involve patients or the public in its design as it is completely desk based. However, our research questions explores improvements in RHI, which are meant to benefit patients and the public accessing health systems.

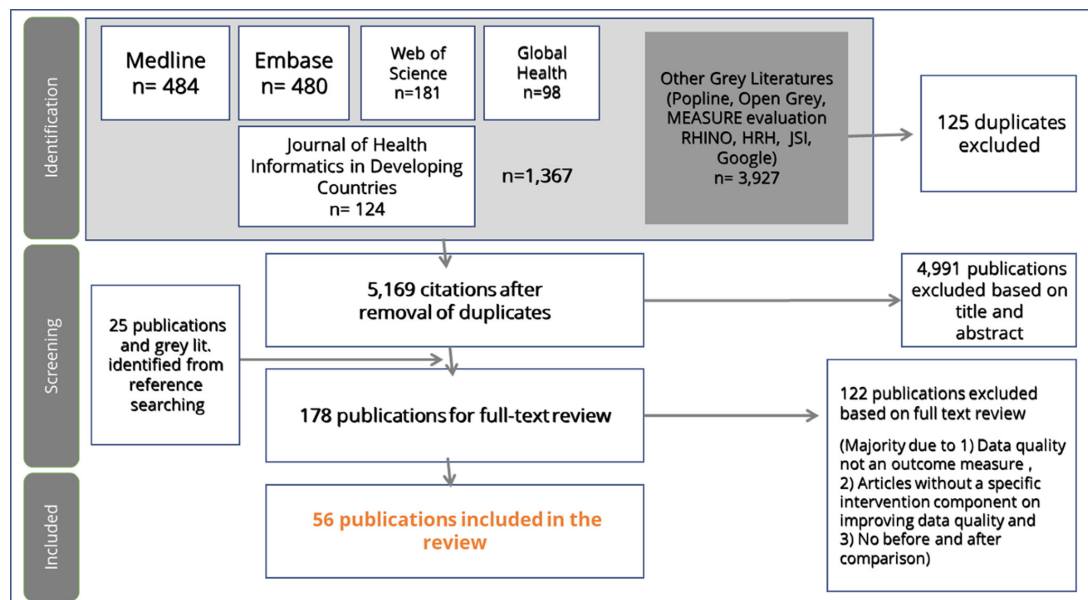


Figure 3 PRISMA flow chart showing the selection of studies. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

RESULTS

Overview of studies included in the review

In total, 5294 references were first identified by the literature search. After removing 125 duplicates, 5169 references were screened. After title and abstract screening, 178 publications were eligible for a full-text review. One hundred and twenty-two publications were excluded because they did not meet our inclusion criteria. Fifty-six peer reviewed articles and published reports were included in this review (figure 3). All studies were published after the year 2000, with most studies (52/56, 93%) published after 2010 (table 1). All but one article was published in English.

The most frequent types of studies included were quasi-experimental (pre–post intervention studies) (n=22, 39%), followed by case studies (n=11, 20%) and cross-sectional studies (n=8, 14%). Five process evaluations (9%), 5 randomised control trials (RCT) (9%), and five retrospective studies (9%) were also included. Most studies were published in peer-reviewed journals (n=45, 80%) and the remaining 11 studies were identified from the grey literature searches. Thirty studies (53%) tested DQ using quantitative methodology, 19 (34%) used mixed-methods and 7 (13%) qualitative methods (table 1).

Studies were carried out in relatively short periods of time, with 24 studies conducted within a study period of <1 year (43%). One-third of studies were conducted over 1–3 years (n=19; 34%) and fewer for more than 3 years (n=10, 18%). Studies were mainly conducted in Africa (n=48, 86%) with most undertaken in Kenya (n=11, 20%). A small number of studies were conducted in Asia (6; 11%), South America (n=3, 5%) and the Pacific Islands (n=1, 2%). Two of the studies described an intervention in a multi-country setting, while the rest were

single country studies (table 1). Figure 2 shows the overview of quantitative and qualitative outcome measures.

A list of selected studies can be found in online supplemental file S1(2.1).

Quantitative outcomes

Out of the studies that measured quantitative outcomes (n=44), more than half of studies measured more than one DQ output (ie accuracy, completeness, timeliness and relevance) (n=26, 59%). There were ten studies (23%) that measured more than three outputs; accuracy, completeness and timeliness together. The most common output measured was completeness (n=34, 77%), followed by accuracy (n=29, 66%), timeliness (n=15, 34%) and relevance (n=1, 2%). DQ output metric (accuracy, completeness, timeliness) was defined in diverse ways in the selected studies. For example, in separate studies Alhanhazo *et al.* define completeness indistinctly in 2014 as ‘exhaustiveness’²² and then as ‘no missing data’²³ in 2015. Whereas, Kintu *et al.*²⁴ describe it as ‘The proportion of health facilities reporting out of the total number of units in the districts’ (Examples of DQ output definitions by included study is available in online supplemental file S1(2.4)). Overall, majority of (n=41, 93%) studies reported improvement in one or more DQ outputs. Studies that tested completeness showed highest proportion of outcomes with improvement (29/34, 85%), followed by timeliness (n=12/15, 80%) and accuracy (23/29, 79%). Only one study measured data relevance and showed improvement (table 2).

When the ‘above threshold’ standard (outcome improved and ≥80% postintervention) was applied, 30 (68%) studies reported one or more outputs with improvement that resulted above threshold (table 2), which is 25% lower than proportion of studies that showed

Table 1 Summary of selected studies (n=56)

No of studies (%)			
Publication date		Study setting*	
Before 2010	4 (7%)	Africa	48 (86%)
After 2010	52 (93%)	Asia	6 (11%)
		South America	3 (5%)
		Pacific Islands	1 (2%)
Type of publication			
Peer-reviewed articles	45 (80%)		
Grey literature	11 (20%)		
		Methodology	
Study design		Quantitative study	30 (54%)
Quasi-experimental studies	22 (39%)	Qualitative study	7 (13%)
(Pre–postintervention)		Mixed-methods study	19 (34%)
Case studies	11 (20%)		
Cross-sectional studies	8 (14%)	RHIS factors addressed by the interventions	
Process evaluations	5 (9%)	All factors addressed	43 (77%)
Randomised controlled trials	5 (9%)	Technical factors addressed	49 (88%)
Retrospective studies/non-experimental design	5 (9%)	Organisational/ Behavioural factors addressed	51 (91%)
Study period		RHIS processes addressed by the interventions	
<1 year	24 (43%)	Multiple processes addressed	48 (86%)
Between 1 and 3 years	19 (34%)	Most addressed process	Data collection (52;93%)
More than 3 years	10 (18%)	Least addressed process	Data display (8;14%)
Unknown	3 (5%)		

*Includes two multi-country study: one with multiple African countries and one with countries in both Africa and Asia
RHIS, routine health information system.

improvement in DQ. After studies without an outcome measure expressed in per cent was removed, studies that tested completeness showed highest proportion of above threshold outcomes (19/31, 61%), followed by accuracy (14/23, 60%), and timeliness (n=3/12, 25%).

Summary of outcomes of each selected study can be found in online supplemental file S1 (2.1).

Qualitative outcomes

In total, 14 studies included qualitative outcomes. Four studies used only qualitative methods and eight studies used mixed methods (in total there were ten mixed-method papers but only six had qualitative outcomes). Qualitative studies aimed to understand the perceptions of users of DQ interventions; particularly whether they perceived the interventions to be useful in improving DQ. Mixed-methods papers often looked at outcomes of behavioural interventions such as trainings alongside process strengthening interventions such as mobile phone data collection. Qualitative studies also attempted to better understand what factors facilitated the creation of an enabling environment. For example, Measure Evaluation (2019)²⁵ describes technical factors such as standardisation of tools, but also a ‘leading from the behind’ approach that they attribute to building a sense of ownership of the health information system as a whole.

Studies that captured qualitative outcomes focused on end-user perception of the intervention. For example, several studies cited variability of participant perceptions of electronic interventions; this included agreement that while some interventions could improve data accuracy there were doubts that indicators that were contingent on human behaviours (ie, timeliness and completeness) could similarly improve.

These studies often attempted to understand user perceptions of improvements to the key data outputs variables. Key challenges described in qualitative research included: difficulties fully defragmenting HIS, where parallel data collection systems continued despite interventions to streamline data collection. The continued use of parallel systems was described in one instance as a necessity brought about by ‘data demands to show the success of programmatic investments year by year’²⁶ to which the national data system could not respond. Additional challenges include limited infrastructure that prevented complete coverage with e-systems, interoperability, maintenance of DQ achievements and evidence use at all levels of the health system.

DQ intervention components

Seventeen different intervention components (ie, discrete activities within an intervention) were

Table 2 Summary of outcomes for studies with quantitative outcomes (n=42)

	No of studies (%)
Data quality outputs measured	
More than one data quality outputs measured	26 (59)
Data accuracy measured	29 (66)
Data completeness measured	34 (77)
Data timeliness measured	15 (34)
Data relevance measured	1 (2)
Improvement in data quality output	
One or more outputs improved	41/44 (93)
Improvement in data accuracy	23/29 (79)
Improvement in data completeness	29/34 (85)
Improvement in data timelines	12/15 (80)
Improvement in data relevance	1/1 (100)
Data quality above threshold (outcome improved and ≥80% after intervention)	
One or more outputs with adequate data quality after improvement	30/44 (68)
Improved and adequate data quality in accuracy	14/23 (60)
Improved and adequate data quality in completeness	19/31 (61)
Improved and adequate data quality in timeliness	3/12 (25)
Improved and adequate data quality in relevance	1/1 (100)
Improved and adequate data quality in general	N/A

N/A, not available.

identified in selected studies (figure 4). Seven were technical (ie improving paper-based data collection forms, using mobile-Health (mHealth) solutions, eHMIS, equipment purchase and maintenance, conducting DQA, improving data storage, and database harmonisation). Ten were organisational/behavioural (ie training, enhanced training, task-shifting and creation of new roles, supervision, enhanced supervision, engagement of core partners in the intervention, dissemination meetings, incentives and standardised protocols), and one was both technical and organisational/behavioural (ie. DQ checking). Most studies (n=43, 77%) included both the technical and organisational/behavioural component in their interventions.

RHIS processes addressed in reviewed studies

Most studies (n=48, 86%) addressed multiple processes of the RHIS, and over two-thirds addressed more than three processes (n=33, 59%). Many of the interventions addressed the data collection stage (n=51, 91%). Data transmission and processing stages were often addressed together when an intervention aimed to improve the data collection stage, since there were numerous eHMIS systems establishment such as the widely used District

Health Information Software 2 (DHIS 2) or mHealth interventions such as Short Message Service (SMS) or mobile applications. DQ checking (n=27, 48%) and feedback stages (n=21, 38%) were addressed with interventions such as enhanced supervision, routine DQ checking by supervisors or automatic DQ check using in-built computer software. Data analysis (n=18, 32%) and display (n=9, 16%) stages were the least frequently addressed stages. A complete list of processes addressed per selected studies can be found in online supplemental file S1(2.2).

COMPARISON OF INTERVENTION COMPONENTS IMPLEMENTED IN ABOVE AND BELOW THRESHOLD STUDIES FOR EACH DQ OUTPUTS

Accuracy

We identified intervention components common to quantitative studies that defined as ‘above threshold’.

Figure 5 shows the interventions components implemented in studies that were greater than or equal to 80% accuracy after implementation compared with those that were not. These are ranked by the greatest difference in the % of studies that assessed those intervention components. Ten intervention components were most frequently deployed in studies with above threshold results on DQ, these were (in order of frequency); training (11 vs 3), equipment (4 vs 0), DQ checks (6 vs 2), improved paper-based tools (3 vs 0), task-shifting (2 vs 0), meetings (5 vs 2), eHMIS (5 vs 2), engagement of stakeholders (1 vs 0) and enhanced supervision (1 vs 0). Above-threshold studies also tended to have implemented more intervention components

Similar comparisons were made for RHIS processes addressed by interventions in studies that measured change in data accuracy (online supplemental file S1(2.5)). Four RHIS processes (data processing, DQ checking, feedback and data display) was most frequently used in studies effective in improving data accuracy. In addition, studies that addressed more than three processes were more effective in improving DQ.

Completeness

Nine intervention components were implemented more in above threshold studies, these included (in order of greater difference) meetings (4 vs 1), engagement of stakeholders (7 vs 3), eHMIS (10 vs 5), task-shifting (2 vs 0), enhanced supervision (3 vs 1), enhanced training (3 vs 1), training (14 vs 8), equipment (2 vs 1) and mHealth (5 vs 3). Above-threshold studies also tended to have implemented more intervention components (figure 6).

Most frequently addressed RHIS processes in above threshold studies were data processing, data transmission, DQ checking, and data analysis, in order of sequence according to larger difference, and those with more than three processes addressed tended to be more effective (online supplemental file S1(2.6)).

Timeliness

Only three intervention components produced outcomes above-threshold among studies aimed at improving

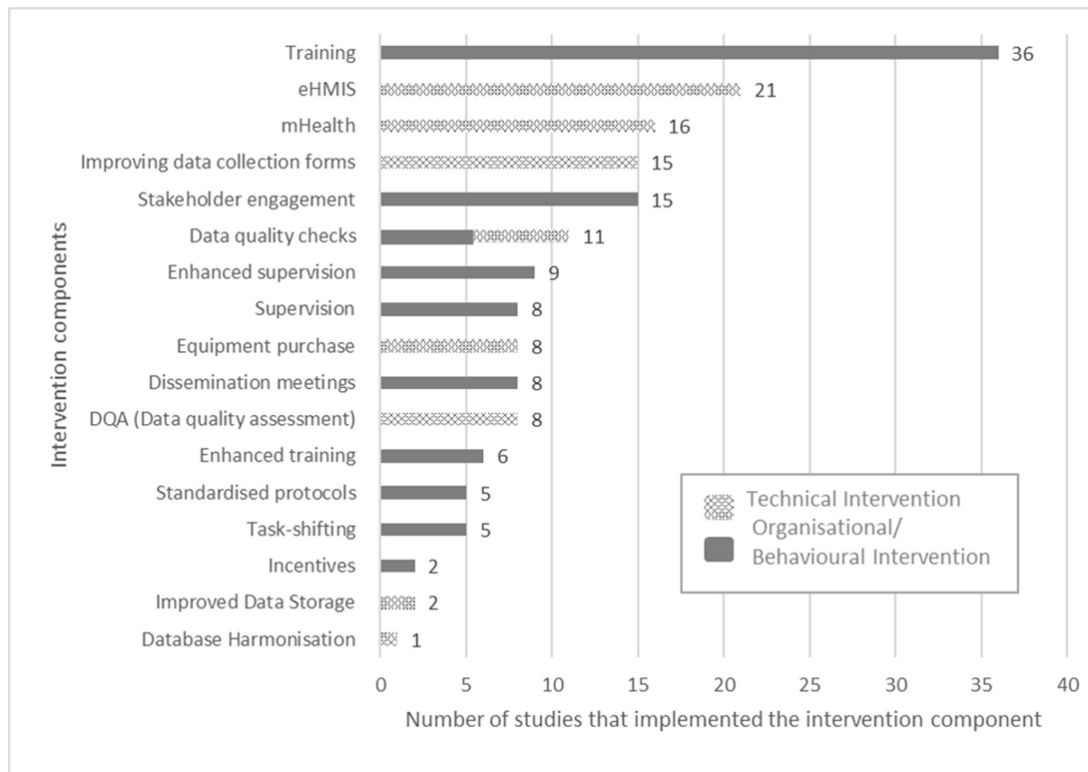


Figure 4 RHIS data quality intervention components implemented in selected studies. eHMIS, electronic health management information system; RHIS, routine health information system.

timeliness (n=7). These were mHealth (2 vs 1), enhanced training (2 vs 1), and eHMIS (2 vs 4). Above-threshold studies also tended to have implemented more intervention components. As for the RHIS processes addressed, feedback process was the only component that was addressed most frequently in above-threshold studies.

DISCUSSION

Using the PRISM framework, we examined components of interventions designed to increase the quality of data generated through RHISs at peripheral levels of the

health system. We screened 5294 references and identified 17 interventions in 56 studies designed to increase DQ and that typically measured improvements of three outputs: accuracy, completeness and timeliness—although definitions of those outputs differed between studies (online supplemental file S1 (2.4)). Most studies were conducted in Africa, and in particular Kenya (20%). This may reflect a concentration of research and NGO (Non-Governmental Organisation) communities located in Africa and Kenya working in the field of DQ.

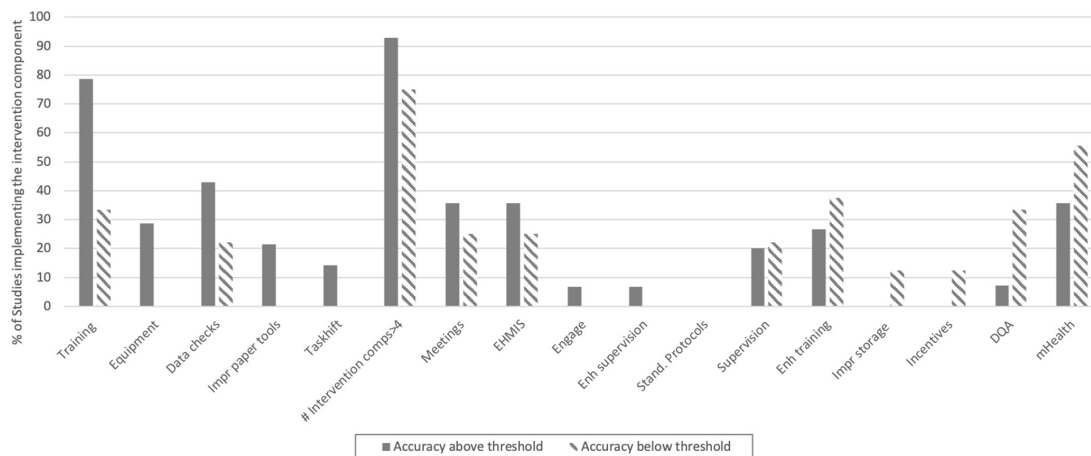


Figure 5 Intervention components in studies that were effective in improving data accuracy above threshold (≥80%) versus that did not improve or did not reach the threshold, ranked by the greatest difference in the percentage of studies with intervention components. impr, improved; Stand, standardised, Enh, enhanced.

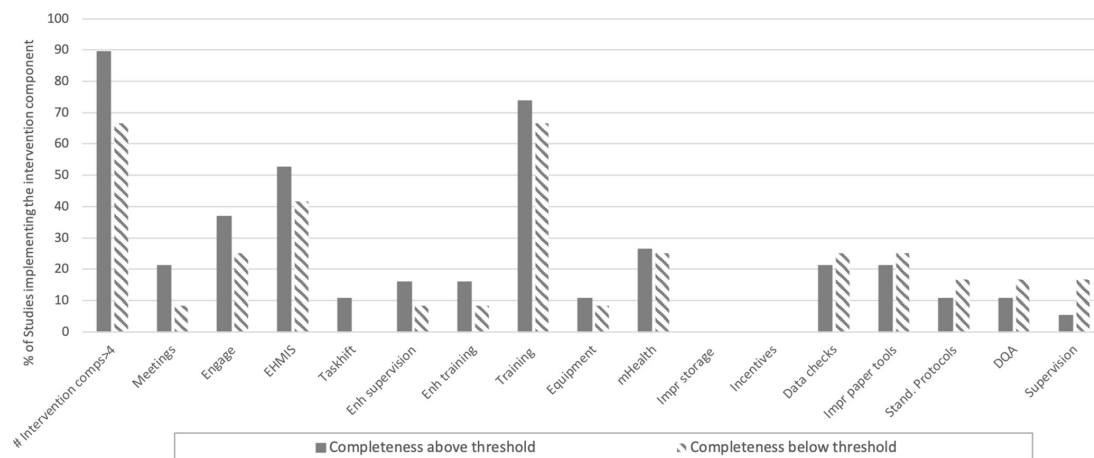


Figure 6 Intervention components in studies that were effective in improving data completeness above threshold ($\geq 80\%$) versus that did not improve or did not reach the threshold, ranked by the greatest difference in the percentage of studies with intervention components. impr, improved; Stand, standardised, Enh, enhanced. DQA, data quality assessment; eHMIS, electronic health management information system.

Most studies (93%) reported improvements in DQ, regardless of the types of interventions or context. This suggests that any investment in improving data accuracy, completeness and timeliness may result in DQ improvements, though not necessarily to a satisfactory level of quality. We attempted to identify more effective interventions by setting a DQ threshold of 80% for any of the DQ outputs measured. Intervention components that improved all three DQ outputs from below to above the 80% threshold were training (either normal or enhanced) and eHMIS. Examining DQ outputs separately, completeness was improved most through meetings, engagement of stakeholders and eHMIS. Accuracy was improved through training, equipment purchases and maintenance, data checks and timeliness through mHealth, enhanced training and eHMIS.

Overall, eHMIS was the intervention that was most frequently associated with increased DQ above the 80% threshold. DHIS2 was the most used eHMIS system. Training was the second most used intervention that seemed to effectively improved DQ in terms of accuracy and completeness (figures 5 and 6). The subject of the training varied and included data aggregation,²⁷ new data entry form²⁸ and utilisation of new digital systems.^{29–45} The effect of enhanced training with longer training period did not seem to exceed the effect of short-term training in improving accuracy and completeness (figure 5 and 6), this is consistent with findings by Rowe *et al.*⁴⁶ The use of mHealth solutions alone was only effective in improving timeliness, using mHealth interventions in reducing the time lag between collection and usage points had an impact on improving timeliness. Conversely, it did not show an effect in preventing input errors at the point of collection and thus was not enough to ensure data accuracy.^{37 47 48} For this reason, authors of the studies included in this review recommended that in addition to an automated system at the point of collection, DQ checking and supervision in multiple levels

of the system is crucial to ensure better accuracy and completeness.^{36 49–56}

Overall, technical interventions alone were not shown to be ‘silver-bullets’, but required careful consideration of context.^{9 57 58} For example, eHMIS implemented as part of the intervention in most studies, consisted of health system-wide components that addressed multiple processes from data collection through to feedback. Indeed, while we have described single components most frequently associated with effective DQ improvements, our findings also suggest that more comprehensive approaches in the design of DQ interventions, that is, applying more intervention components and addressing all the technical, organisational, and behavioural aspects of RHIS, were likely lead to greater improvements in DQ than implementing a single component interventions (figures 5 and 6). Organisational/behavioural factors seem particularly important given that all but one of the top five most effective interventions that improved data completeness included interventions that included these such as meetings, engagement of stakeholder, task-shifting, enhanced supervision. This finding of the need for a more holistic approach is not particularly novel. With the advent of the ‘microcomputer’ in the 90s, Sandiford *et al* had already identified that technical approaches alone would not improve DQ.⁹

Different DQ outputs are interlinked (ie, accuracy without completeness cannot exist), but this review showed that the mechanisms of improving each output may differ. Pilot, pre–postevaluations or controlled trials provide important insights into elements that are likely to impact on primary outcomes, however, beyond the study period these outcomes must be regularly monitored when implemented at scale.⁵⁹ Nearly half of the studies (43%) were undertaken over periods of less than 12 months. One might expect optimal results during intense investigate periods. The frequency of studies of a duration >3 years was relatively low (18%), therefore

limiting our ability to examine sustainability or real-life implementation constraints.

Fewer qualitative (13%, [table 1](#)) and mixed-method studies (34%, [table 1](#)) were identified during the review of the literature. These studies often reported improved DQ based on ‘perceptions’ of the users who were involved in the interventions. Further studies including qualitative methods are necessary to examine appropriateness of interventions or perceived usefulness of different data component applications in various contexts in order to unpack why interventions may not work or could be further optimised.

Our findings show a remarkable diversity in both the methodology to test interventions and the measurement of DQ. Most studies included in the review were pre-postintervention comparisons ([table 1](#)), only five studies were formal RCTs. The latter are considered to provide the highest quality evidence. The design of quality improvement interventions is a critical area for future studies focusing on RHIS.⁶⁰

In terms of measuring DQ outputs, our review highlights the need to design studies of RHIS interventions with a clear set of measurable outputs, which are comparable beyond a trial or pilot phase. While we defined ‘above-threshold’ studies as those that showed improvement and achieved $\geq 80\%$ DQ postintervention for any of the three DQ outputs, a more comprehensive approach would have been to apply thresholds set out in the DQ review⁶¹ in which multiple thresholds are suggested according to different levels of the health system and core indicators for health data used universally (ie, Antenatal Carefirst visit, third-dose DTP (diphtheria, tetanus, pertussis)-containing vaccine). Going forward, studies investigating the effect of interventions on DQ should aim to align evaluations with thresholds and targets set out in recent years.

Limitations across studies

Most studies provided limited details on their interventions and this could have missed components. We relied on authors to report key intervention activities, which could introduce an inherent bias. Also, some components might appear more effective, due to the number of studies implementing that component. For example, components that were included in less studies, such as ‘task-shifting’, which was only in five studies, appeared more effective in improving completeness as defined as the difference between studies that showed improvement vs those that did not. Comparatively, training was included in many studies and while still ranked as a top five component appears ‘less effective’ than task shifting. It is still important to note that within our analysis the top three to five components showed a clear difference.

While cost-effectiveness was not an outcome measure for this study, we note that only three of the studies^{34 62 63} considered costs such as cost–benefit analysis of staff time or comparison of cost between different digital interventions, but none assessed costs per quantity of data

improvement and this is something often ignored. This omission was certainly a limitation across studies and prevented the authors’ ability to assess intervention components based on opportunity costs and budget allowances.

Limitations of this review

There are several limitations to this review related to the number of studies that we identified and were able to include into the study and the variability in measurement of three DQ outputs. First, we did not conduct a meta-analysis in this review due to varying outcome measures for each of the DQ outputs (ie, accuracy, completeness, timeliness, relevance) which could not be consolidated (online supplemental file S1(2.4)), and were not clinical data. Furthermore, only five RCTs were identified, making it unviable to consolidate quantitative measures to conduct a meta-analysis.

Throughout the studies, the number of times an intervention component was studied varied as did the number of DQ outputs that studies measured. This limited the analyses that we could undertake and the strength of the conclusions we could draw from the review. In examining DQ outputs in an aggregated way, we risk being unable to disentangle which interventions contributed to which DQ output measures. However, disaggregating studies by DQ output measured means reviewing small numbers of studies. We tried to strengthen our analyses by both aggregating and then disaggregating DQ outputs measured to assess whether there were key intervention components emerging as key to improving overall DQ. Additionally, due to length considerations we have presumed that the interventions, distinguished across 17 different categories, were implemented with similar intensity across contextually similar settings, which is not reflective of reality.

More broadly, our study did not address data use—which is likely to act as an intervention in and of itself because the use of data leads to feedback and several studies, including this review, have shown that feedback does improve DQ.

Finally, the interventions we identified do not address some fundamental factors identified as challenging to good DQ previously¹⁷ such as ‘technical infrastructure, issues such as unreliable electric power and erratic Internet connectivity and clinicians’ limited computer skills... good communication and networking actions among all stakeholders of HMIS, and information culture at different levels of district health information systems’.⁶⁴ Given studies included in this review tended to take place over relatively short periods of time, these more fundamental issues may not have been identified as barriers to uptake of the intervention more widely.

CONCLUSIONS

Poor RHI data hamper a country’s efforts to effectively monitor health programmes.^{65 66} Challenges associated

with RHIS, often result on a dependence on optimistic, infrequent sampled household and community surveys for health data.⁷ Efforts are required to improve the collection of quality assured, timely routine data,¹³ managed by new innovative interventions that would build confidence in the fidelity of real-time, nationally owned, routine data.

Holistic DQ interventions addressing not only one determinant (ie, technical) but multiple aspects of the RHIS have shown to be more effective and should be a continued practice. Continued attention and investment on RHI DQ, regardless of type of interventions used, is likely to enable further improvement in RHIS DQ in the LMICs. However, the measures of DQ evaluation standards, and the expectations of what constitutes an ‘improvement’ in DQ should be better defined. Mechanisms to improve data accuracy, completeness and timeliness may differ, thus interventions should be based on thorough assessment and careful design according to the aim of the intervention. A more comprehensive and rigorous evidence platform is still required to provide the basis for future appropriate mixes of effective interventions in LMICs, using a standardised methodology and outcome definitions.

Author affiliations

¹Department of Disease Control, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, UK

²Policy and Programmes Division, World Vision UK, Milton Keynes, UK

³Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK

⁴Population and Health Unit, KEMRI - Wellcome Trust Research Programme, Nairobi, Kenya

⁵Nuffield Department of Clinical Medicine, University of Oxford Centre for Tropical Medicine and Global Health, Oxford, UK

⁶Medical and Healthcare Innovation, British Heart Foundation, London, UK

⁷Department of Health Policy, London School of Economics and Political Science, London, UK

⁸Centre for Universal Health, Global Health Programme, Chatham House, London, UK

Twitter Justin Parkhurst @justinparkhurst

Contributors JL conducted the study and drafted the manuscript; NAE supervised the study and guided the writing of the manuscript and was part of the study selection of studies between October 2017 and January 2021; CAL conceived the paper, supported the data analysis and supported the drafting of the manuscript and was part of the study selection of studies between October 2017 and January 2021; LOH was the second reviewer during the study selection; RWS supported the drafting of the discussions section; NDH was part of the study selection of studies in January 2021, JP and JW reviewed and revised the manuscript. All authors read, reviewed and approved the final manuscript.

Funding Support for this work was provided by the UK’s Department for International Development (DFID) for their continued support to a project Strengthening the Use of Data for Malaria Decision Making in Africa (DFID Programme Code 203155) that provided support to JL, CAL, LOH and NAE. RWS is funded by Wellcome Trust Principal Fellowship (number 103602 and 212176) and acknowledges the support of the Wellcome Trust to the Kenya Major Overseas Programme (number 203077). This research is also supported by funding from MSD, through its MSD for Mothers programme. MSD has no role in the design, collection, analysis, and interpretation of data, in the writing of manuscripts, or in decisions to submit manuscripts for publication. The content of all publications is solely the responsibility of the authors and does not represent the official views of MSD. MSD for Mothers is an initiative of Merck & Co., Inc., Kenilworth, N.J., U.S.A

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on request. All data relevant to the study are included in the article or uploaded as online supplemental information. The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Justin Parkhurst <http://orcid.org/0000-0003-0831-6213>

Ngozi A Erundu <http://orcid.org/0000-0003-0268-6838>

REFERENCES

- 1 WHO. Everybody’s business: strengthening health systems to improve health outcomes: WHO’s framework for action. *Production* 2007;1:1–56.
- 2 Leon N, Brady L, Kwamie A, *et al*. Routine health information system (RHIS) interventions to improve health systems management. *Cochrane Database Syst Rev* 2015;326.
- 3 Mutale W, Chintu N, Amoroso C, *et al*. Improving health information systems for decision making across five sub-Saharan African countries: implementation strategies from the African health Initiative. *BMC Health Serv Res* 2013;13:S9.
- 4 AbouZahr C, Boerma T, Hogan D. Global estimates of country health indicators: useful, unnecessary, inevitable? *Glob Health Action* 2017;10:1290370.
- 5 Stevens GA, Alkema L, Black RE, *et al*. Guidelines for accurate and transparent health estimates reporting: the gather statement. *Lancet* 2016;388:e19–23.
- 6 Byass P, de Courten M, Graham WJ, *et al*. Reflections on the global burden of disease 2010 estimates. *PLoS Med* 2013;10:e1001477.
- 7 Alegana VA, Okiro EA, Snow RW. *Routine data for malaria morbidity estimation in Africa: challenges and prospects*, 2020: 1–13.
- 8 Hotchkiss DR, Diana ML, Foreit KGF. How can routine health information systems improve health systems functioning in low-and middle-income countries? Assessing the evidence base. In: *Health information technology in the International context*. 44. Emerald Group Publishing Limited, 2012.
- 9 Sandiford P, Annett H, Cibulskis R. What can information systems do for primary health care? an international perspective. *Soc Sci Med* 1992;34:1077–87.
- 10 Lippeveld T, Sauerborn R, Bodart C. *Design and implementation of health information systems*. 280. WHO, 2000.
- 11 Braa J. Health information systems in developing countries. *Vital wave Consult* 2009;69.
- 12 AbouZahr C, Boerma T. Health information systems: the foundations of public health. *Bull World Health Organ* 2005;83:578–83.
- 13 Chan M, Kazatchkine M, Lob-Levy J, *et al*. Meeting the demand for results and accountability: a call for action on health data from eight global health agencies. *PLoS Med* 2010;7:e1000223–8.
- 14 Health Data Collaborative. *Who we are - health data collaborative*, 2020.
- 15 DeLone WH, McLean ER. Information systems success: the quest for the dependent variable. *Inf Syst Res* 1992;3:60–95.
- 16 Aqil A, Lippeveld T, Hozumi D. Prism framework: a paradigm shift for designing, strengthening and evaluating routine health information systems. *Health Policy Plan* 2009;24:217–28.
- 17 Ndabarora E, Chipps JA, Uys L. Systematic review of health data quality management and best practices at community and district levels in LMIC. *Inf Dev* 2014;30:103–20.

- 18 Aqil A. *Using PRISM to strengthen and evaluate health information systems*. Chapel Hill, North Carolina: University of North Carolina at Chapel Hill, Carolina Population Center [CPC], MEASURE Evaluation, 2007.
- 19 Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6:e1000097.
- 20 The World Bank. *World bank country and lending groups*, 2017.
- 21 Fall IS, Rajatonirina S, Yahaya AA, et al. Integrated disease surveillance and response (IDSR) strategy: current status, challenges and perspectives for the future in Africa. *BMJ Glob Health* 2019;4:1–11.
- 22 Ahanhanzo YG, Kpozèhouen A. Community workers contribution for data quality improvement in routine health information system: live births number estimation by capture recapture method. *Open J Epidemiol* 2014.
- 23 Ahanhanzo YG, Saizonou J, Wodon A. [Health workers' involvement for data quality improvement in Benin]. Implication des agents de santé dans la conception des outils de collecte de données au Bénin. *Sante Publique*;27:241–8.
- 24 Kintu P, Nayunja M, Nzabanita A. Development of HMIS in least developed country settings: the case for Uganda. *IFHRO Congr {&} AHIMA Conv* 2004:1–9.
- 25 MEASURE Evaluation. *Lessons in health information system strengthening what worked in Mali lessons in health information system strengthening what worked in Mali*, 2019.
- 26 MEASURE evaluation. *Strengthening Côte d'Ivoire's health information system to combat HIV: validating data on the effectiveness of interventions*, 2018.
- 27 Admon AJ, Bazile J, Makungwa H, et al. Assessing and improving data quality from community health workers: a successful intervention in Neno, Malawi. *Public Heal Action* 2013;3:56–9.
- 28 Mwakysua S, Wamae A, Wasunna A, et al. Implementation of a structured paediatric admission record for district hospitals in Kenya – results of a pilot study. *BMC Int Health Hum Rights* 2006;6:9.
- 29 Amoroso CL, Akimana B, Wise B, et al. Using electronic medical records for HIV care in rural Rwanda. *Stud Health Technol Inform* 2010;160:337–41.
- 30 Auld AF, Wambua N, Onyango J, et al. Piloting the use of personal digital assistants for tuberculosis and human immunodeficiency virus surveillance, Kenya, 2007. *Int J Tuberc Lung Dis* 2010;14:1140–6.
- 31 Ndira SP, Rosenberger KD, Wetter T. Assessment of data quality of and staff satisfaction with an electronic health record system in a developing country (Uganda): a qualitative and quantitative comparative study. *Methods Inf Med* 2008;47:489–98.
- 32 Pascoe L, Lungo J, Kaasbøll J. *Collecting integrated disease surveillance and response data through mobile phones*, 2012: 2–6.
- 33 Rosewell A, Ropa B, Randall H, et al. Mobile phone-based syndromic surveillance system, Papua New Guinea. *Emerg Infect Dis* 2013;19:1811–8.
- 34 Were MC, Kariuki J, Chepng'eno V, et al. Leapfrogging paper-based records using handheld technology: experience from Western Kenya. *Stud Health Technol Inform* 2010;160:525–9.
- 35 Yugi J, Buesseler HM. Text messaging APP improves disease surveillance in rural South Sudan. *Mhealth* 2016;2:8.
- 36 Huang F, Cheng S, Du X, et al. Electronic recording and reporting system for tuberculosis in China: experience and opportunities. *J Am Med Inform Assoc* 2014;21:938–41.
- 37 Tobgay T, Samdrup P, Jamtsho T, et al. Performance and user acceptance of the Bhutan febrile and malaria information system: report from a pilot study. *Malar J* 2016;15:52.
- 38 Barnett D, Yosellina SS, et al. *Mixed-method impact evaluation of a mobile phone application for nutrition monitoring in Indonesia*. Brighton, United Kingdom: Institute for Development Studies [IDS], 2016.
- 39 Gilbert SS, Thakare N, Ramanujapuram A, et al. Assessing stability and performance of a digitally enabled supply chain: retrospective of a pilot in Uttar Pradesh, India. *Vaccine* 2017;35:2203–8.
- 40 Gold J, Andrews H, Appleford G. Using mobile phone text messages (SMS) to collect health service data: lessons from social franchises in Kenya, Madagascar and the Philippines. *J Health Inform Dev Ctries* 2012;6:467–79.
- 41 Haskeew J, Ro G, Saito K, et al. Implementation of a cloud-based electronic medical record for maternal and child health in rural Kenya. *Int J Med Inform* 2015;84:349–54.
- 42 Ishijima H, Mapunda M, Mndeme M, et al. Challenges and opportunities for effective adoption of HRH information systems in developing countries: national rollout of HRHIS and TIIS in Tanzania. *Hum Resour Health* 2015;13:48.
- 43 Kagiri M, Waiganjo P, Ngechu R. *Enhancing community based health information system (CBHIS) reporting through open-source short message service-based tool*. , 2015: 3, 475–84.
- 44 Kariuki JM, Manders E-J, Richards J, et al. Automating indicator data reporting from health facility EMR to a national aggregate data system in Kenya: an Interoperability field-test using OpenMRS and DHIS2. *Online J Public Health Inform* 2016;8:e188.
- 45 Medhanyie AA, Spigt M, Yeboyo H, et al. Quality of routine health data collected by health workers using smartphone at primary health care in Ethiopia. *Int J Med Inform* 2017;101:9–14.
- 46 Rowe AK, Kachur SP, Yoon SS, et al. Caution is required when using health facility-based data to evaluate the health impact of malaria control efforts in Africa. *Malar J* 2009;8:209.
- 47 De TC, Unfried K. *Evaluation of a mobile reporting system for the collection of routine malaria data in Mali*, 2014.
- 48 Medhanyie AA, Moser A, Spigt M, et al. Mobile health data collection at primary health care in Ethiopia: a feasible challenge. *J Clin Epidemiol* 2015;68:80–6.
- 49 USAID, PEPFAR, MEASURE Evaluation. *A case study to measure national HIV M{&}E system strengthening Nigeria*, 2014.
- 50 Wagenaar BH, Gimbel S, Hoek R, et al. Effects of a health information system data quality intervention on concordance in Mozambique: time-series analyses from 2009–2012. *Popul Health Metr* 2015;13:9.
- 51 Chisha Z, Larsen DA, Burns M, et al. Enhanced surveillance and data feedback loop associated with improved malaria data in Lusaka, Zambia. *Malar J* 2015;14:222.
- 52 Mphatswe W, Mate KS, Bennett B, et al. Improving public health information: a data quality intervention in KwaZulu-Natal, South Africa. *Bull World Health Organ* 2012;90:176–82.
- 53 Harrison K, Rahimi N, Danovaro-Holliday MC, Carolina Danivaro-Holliday M. Factors limiting data quality in the expanded programme on immunization in low and middle-income countries: a scoping review. *Vaccine* 2020;38:4652–63.
- 54 Hagel C, Paton C, Mbevi G, et al. Data for tracking SDGs: challenges in capturing neonatal data from hospitals in Kenya. *BMJ Glob Health* 2020;5:e002108–10.
- 55 Wandera SO, Kwagala B, Nankinga O, et al. Facilitators, best practices and barriers to integrating family planning data in Uganda's health management information system. *BMC Health Serv Res* 2019;19:1–13.
- 56 Kagoya HR, Kibuule D. Quality assurance of health management information system in Kayunga district, Uganda. *Afr. eval. j.* 2018;6:1–11.
- 57 Lippeveld T. Routine health information systems: the glue of a unified health system. *Routine Heal Inf Syst Glue a Unified Heal Syst* 2001:14–16.
- 58 Lippeveld T, Sauerborn R, Bodart C. Design and implementation of health information systems, 2000. Available: <https://apps.who.int/iris/handle/10665/42289>
- 59 Wagenaar BH, Hirschhorn LR, Henley C, et al. Data-driven quality improvement in low-and middle-income country health systems: lessons from seven years of implementation experience across Mozambique, Rwanda, and Zambia. *BMC Health Serv Res* 2017;17:830.
- 60 Rowe AK, Rowe SY, Peters DH, et al. Effectiveness of strategies to improve health-care provider practices in low-income and middle-income countries: a systematic review. *Lancet Glob Health* 2018;6:e1163–75.
- 61 WHO. *Data quality review: module 1: framework and metrics*, 2017.
- 62 Kyeyagaire R, Southgate R, Broughton E. *The data management improvement collaborative in Uganda. research and evaluation report*. Bethesda, Maryland: University Research Company [URC], Health Care Improvement Project [HCI], 2011.
- 63 Asangansi I, Macleod B, Meremikwu M. *Improving the routine HMIS in Nigeria through mobile technology for community data collection*. , 2013: 7, 76–87.
- 64 Oluoch T, Santas X, Kwaro D, et al. The effect of electronic medical record-based clinical decision support on HIV care in resource-constrained settings: a systematic review. *Int J Med Inform* 2012;81:e83–92.
- 65 Ronveaux O, Rickert D, Hadler S, et al. The immunization data quality audit: verifying the quality and consistency of immunization monitoring systems. *Bull World Health Organ* 2005;83:503–10.
- 66 Wetherill O, Lee C, Dietz V. Root causes of poor immunisation data quality and proven interventions: a systematic literature review. *Ann Infect Dis Epidemiol* 2017;2:1–7.