


Data Equity to Advance Health and Health Equity in Low- and Middle-Income Countries: A Scoping Review

Digital Health
Volume 7: 1–17
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20552076211061922
journals.sagepub.com/home/dhj


So O'Neil¹ , Sydney Taylor²  and Anitha Sivasankaran³ 

Abstract

Objective: To assess a common hypothesis that data serve as a mechanism to improve health and health equity in low-and middle-income countries (LMICs), we conducted a synthesis of the evidence about the linkage between data capabilities in LMICs and health outcomes.

Methods: We searched and reviewed peer-reviewed and grey literature published in the past decade that focused on at least one aspect of health data or health equity or provided insights on the relationship between data use and improved health outcomes, decision-making, or both. We supplemented this with expert interviews and convenience-sampled literature.

Results: Of the 50 included articles, 33 discussed data collection, with 23 stating that poor accuracy, reliability, and completeness hindered data-informed decision-making. Of 27 articles discussing data access, 18 described how lack of interoperability between data systems hampered governments' and other organizations' ability to leverage the full value of data available. Of 19 articles discussing data use, 13 discussed how data were not getting to those doing work on the ground. Although key informants postulated a virtuous cycle between data and improved health outcomes, evidence did not support this connection.

Conclusions: Findings indicate better data might improve health service delivery. However, more work is needed to examine whether improvements in data yield improvements in health outcomes in LMICs. Our conceptual framework of data equity for health and health equity developed through this scoping review helps identify the key components along which to assess improvements in LMICs' data capabilities.

Keywords

Data equity, health equity, health outcomes, conceptual framework, low- and middle-income countries, data collection, data access, data use, data-driven decision-making

Submission date: 14 May 2021; Revision date: 25 October 2021; Acceptance date: 2 November 2021

Introduction

Data could be a core tool for improving health system functioning and population health, especially among low- and middle-income countries (LMICs) that experience poorer population health outcomes. Studies of civil registration and vital statistic systems (CRVS) have shown that well-performing CRVS are consistently correlated with improved health irrespective of a country's wealth, urbanization, and health system coverage.¹ Improvements in

collecting, accessing, and using data in LMICs could represent a key pathway to improve health and achieve health

¹Health Foundations, Mathematica, Cambridge, MA 02139, United States

²Mathematica, Seattle, WA 98121, United States

³Mathematica, Oakland, CA 94612, United States

Corresponding author:

So O'Neil, Health Foundations, Mathematica, 955 Massachusetts Avenue, Suite 801, Cambridge, MA 02139, United States.
Email: soneil@mathematica-mpr.com



equity. Introducing electronic health data systems in LMICs has the potential to improve data quality and efficiency in service delivery. Studies have demonstrated that computerized health information systems also provided small annual cost savings to the public health system.^{2,3} Innovative uses of data outside of the traditional health system have also helped identify health risks and outbreaks before they occur or in real time, allowing health authorities to better coordinate and respond.⁴ However, LMICs continue to have less sophisticated systems for collecting and analyzing data than high-income countries (HICs).^{5–10} In response, the World Health Organization has issued a guide on improving data quality specifically for developing countries.¹¹

Despite the promise of data to improve health, few studies have directly examined the impacts of improved data on health care delivery, health outcomes, and health equity. In addition, the mechanism through which data could potentially improve health remains unclear. Our aim was to conduct a scoping review to assess the current evidence on underlying causes of data inequities for health, the ways in which data can affect health and health equity, and gaps in knowledge about whether and how improving data equity can improve health and health equity. We reviewed evidence on inequities in data collection, access, and use in LMICs and how addressing these aspects of data could improve health and health equity.

Methods

Overview

Our study included a review of academic and grey literature, as well as supplemental interviews with contributors to the field. We conducted our scoping review in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews guidelines.¹²

Search strategy and selection criteria

We searched seven academic literature databases to find relevant articles: Academic Search Premier, Scopus, MEDLINE, PAIS International, Health Policy, Cochrane Database of Systematic Reviews, and PsycINFO. Grey literature search engines included Google Scholar as well as a targeted search of 23 websites (Appendix 1). Our primary search terms included data equity, health equity, low- and middle-income countries (LMICs), health outcomes, community health, digital divide, and data-driven decision-making. Our secondary search terms included 24 health terms, 16 data terms, and 6 population terms combined using Boolean operators. Articles had to be published not more than ten years before the search (January 2009 to July 2019) and available in English. All databases were

searched between July 12, 2019, and July 19, 2019. Our search restrictions may have excluded articles or reports relevant to our review but that did not use our specific search terminology or combination of terms (Appendix 1, Appendix 2). Thus, we examined the reference lists of published review articles to identify additional publications not retrieved through the initial database searches and inquired about relevant articles with lead authors and contributors to the field.

Publication selection

One researcher, ST, screened titles and abstracts. Articles for inclusion met all of the following criteria for review: topical relevance, study setting in a LMIC, peer-reviewed or credible grey literature, and use of scientifically-based research standards according to the Education Science Reform Act of 2002 (Appendix 3). Grey literature was deemed to be credible if the report offered factual, evidentiary support and a review of the abstract or executive summary did not lead to major reservations about the document's partiality or use of scientifically based research standards. Articles were deemed topically relevant if they met at least one of the following criteria: (1) focused on health data, health information systems, data representativeness, data completeness, data gaps, open data, data access, and/or data equity; (2) focused on health outcomes and/or health equity; (3) focused on community health; and (4) insights into the relationship between data use and improved health outcomes and/or improved decision-making. ST uploaded article abstracts into RefWorks for managing the search results and exported the abstracts to Excel to assign articles to reviewers for analysis.

Five researchers (including AS, SO, and ST) divided the screened-in articles and conducted full-text reviews. They abstracted relevant information into a standardized template to document the type of study, whether it was peer reviewed, a description of the sample (if relevant), article setting, a brief design description, key data sources, primary analysis methods, and the likely value to the scoping review (Appendix 4). Articles that were deemed not relevant typically included pilot studies focused on a niche aspect of health, such as studies examining the use of a specific health information system software for a particular disease in one context, or studies that did not offer evidence to support their claims. Researchers categorized abstracted information along the core data dimensions of the scoping review—data collection, data access, and data use—as well as additional categories of interest, such as findings relating health data to improvements in monitoring and accountability, health outcomes, health equity, and opportunities to achieve health or data equity.

The researchers also assessed the level of evidence contained in each publication based on the rigor of the study design and generalizability of findings. Large, multi-site

experimental or quasi-experimental designs (QEDs) or meta-analyses where more than 50% of the included studies were randomized controlled trials (RCTs), QEDs, or longitudinal studies were considered to possess high evidence; studies were considered to be of low evidence if they were single-site, pilot observational studies; pre-post, descriptive, non-peer-reviewed qualitative studies; or position papers that did not present empirical evidence. Studies considered to be of moderate evidence combined elements of high- and low-evidence studies. Objectivity was separately assessed based on the degree to which claims were supported by data and the independence of the research team. All included articles were reviewed twice, with most articles assigned to a different researcher in the second round of review to add a layer of quality assurance. Any discrepancies between the two researchers were discussed and resolved.

Expert interviews

We conducted interviews with contributors to the field to explore prominent themes in greater depth and capture perspectives from practitioners and researchers. We selected these key informants from the literature and organizations included in our targeted grey literature search based on their contributions to current work and research in the field, as well as relevant practical experience to support health data systems in LMICs. We identified 17 key informants for interview. Sixteen were contacted via email; the seventeenth acted as back-up if his or her colleague did not reply. Five of 16 did not reply and were not interviewed. The remaining 11 participated in 30- to 60-min interviews with one researcher leading the interview (AS or SO) and another (ST) taking notes. Among the key informants who were interviewed, 6 were identified through the literature alone, 2 identified through both the literature and funder, and 2 were referred by the funder of the study. One interview was excluded due to relevance; the interviewee was uncooperative and did not directly answer the researchers' questions. Seven key informants represented non-governmental organizations such as PATH and UNICEF, two from philanthropic foundations, and one from an academic institution. All but one of the key informants was based in the United States, with one based in Kenya. All key informants conducted work and research in LMICs. The protocol included questions about participant background and experience related to data for health, perspectives on data disparities between LMICs and HICs, definition of data equity, and evidence on the connection between data equity and health. The protocol also resonance tested ten statements on key findings from the evidence review to assess key informants' agreement with statements and whether they considered the quality and volume of the evidence underlying these messages to be low, medium, or high (Appendix 5). Interviews were

recorded and referenced to supplement any gaps in the notes. ST coded and analyzed the notes using the same coding scheme as that used to abstract information from the evidence review.

Role of funding source

The funder of the study reviewed and provided feedback on the interview questions and the high-level approach of the study. The funder also suggested eight key informants to consider for inclusion in interviews; two were ultimately interviewed. The authors made the final decisions on all aspects of the research, had full access to all the data in the study, and had final responsibility for the final content decision to submit for publication.

Results

The search yielded 405 articles (125 from the database search, 139 from grey literature, and 141 convenience sampled), with 177 reviewed in-depth and 50 (28%) deemed relevant to the themes of the scoping review (Appendix 3); 16 key informants were identified and 11 interviewed (Figure 1). The articles retrieved through our search included 19 studies within a global setting, seven within a multi-country setting, 11 within Sub-Saharan Africa, four within South Asia, and two within Southeast Asia. Theoretical and analytic essays were assigned a global study setting because they typically drew on evidence from many different contexts. Twenty-eight (56%) were peer reviewed, including 15 descriptive studies, five cross-sectional studies, and seven meta-analyses, including three systematic reviews. No study design used an RCT (although articles using meta-analyses may have referenced studies using RCTs) and two were pre-post studies. Many of the articles (58%) employed qualitative methods, another 20% relied on quantitative methods as the primary method analysis, and 22% employed mixed methods (Table 1). Using our criteria based on the rigor of the design and generalizability of findings discussed earlier, most articles (30) were assessed to have low evidence supporting their claims, with four assessed to have no evidence. Twelve articles provided moderate evidence to support their argument, while only four had high evidence.

Data collection for health and health equity

Of the 50 articles, 33 converged around four key factors that influence the quality of data during the collection process and how these factors affect data's potential use for improving health in LMICs. The four key factors included integrity of data recorded, mode of data collection and storage, population sampled, and benefit or burden of conducting data collection (Table 2). Within these same four themes, nine of the ten key informants

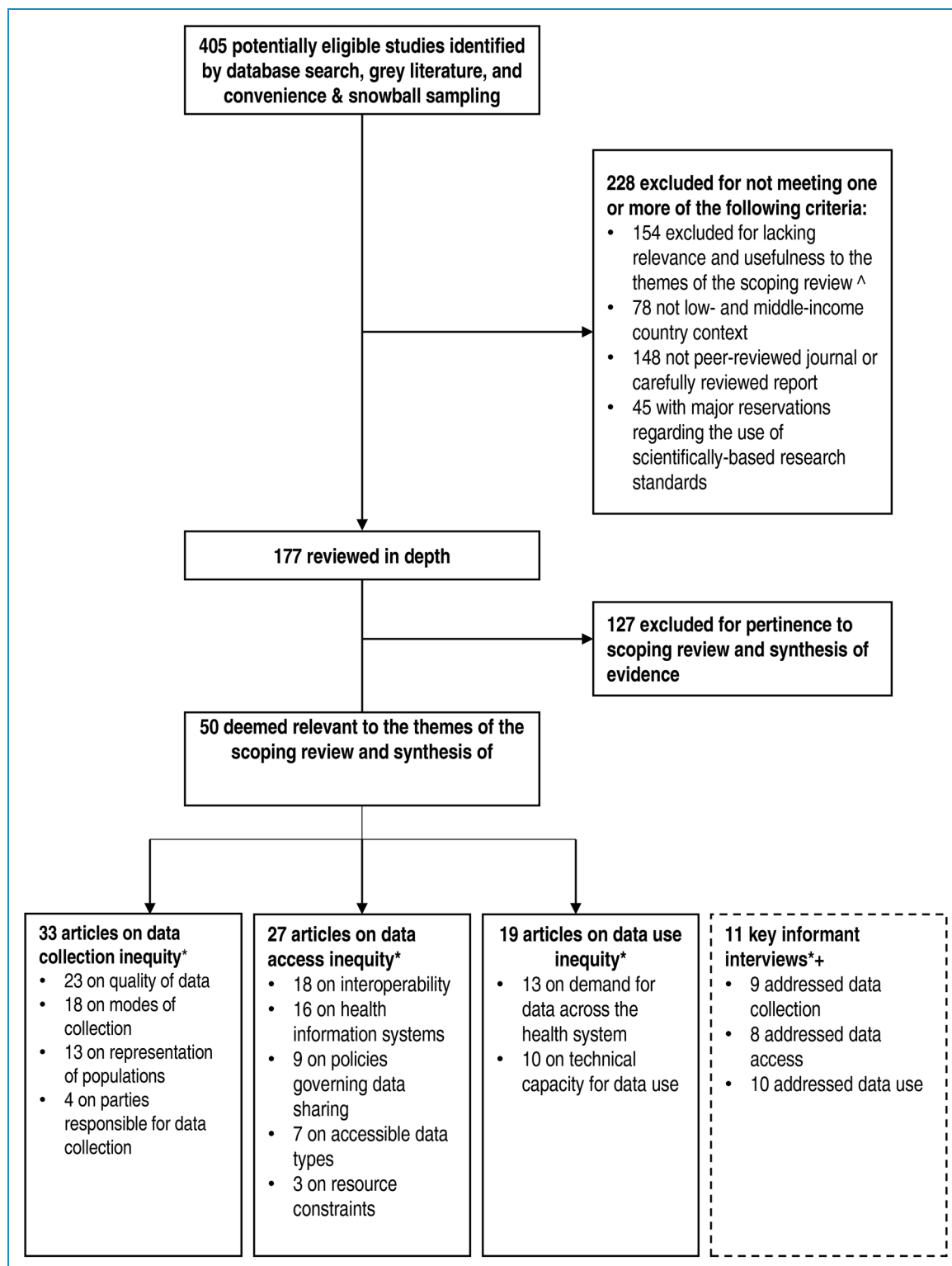


Figure 1. Study selection.

highlighted widespread issues with data collection. Like the articles, the vast majority of key informants (9) most commonly pointed to the integrity of data recorded as a driving factor for data gaps in LMICs. Articles focused

on data collected at one of several levels of the health system, including national, provincial, regional, district, and/or health facility level. More than half (18) discussed biases introduced by mode of data collection.

Table 1. Characteristics of included studies.

Citation	Type of article	Setting	Research method	Topics addressed
Agarwal et al. 2015 ¹³	Systematic review	Global	Qualitative	Data collection, data use, data for health service delivery and resource allocation
Akachi and Kruk 2017 ¹⁴	Analytic essay	Global	Qualitative	Data collection, data use, data for health service delivery and resource allocation
Bhavaraju 2018 ¹⁵	Editorial	India	Qualitative	Data collection, data access
Burke 2013 ¹⁶	Analytic essay	Global	Qualitative	Data use
Buvinic et al. 2014 ¹⁷	Report	Global	Qualitative	Data collection, data access, health equity
Carney 2015 ¹⁸	Analytic essay	West Africa/ US	Qualitative	Data access, data use, data for health service delivery and resource allocation
Chandy and Zhang 2015 ¹⁹	Analytic essay	Global	Qualitative	Data collection
Devex, Philips 2019 ²⁰	Report	Multi-country	Qualitative	Data access, data use
Endriyas et al. 2019 ²¹	Research article	Ethiopia	Quantitative	Data collection, data access, data use, data monitoring and accountability of the health system
Flahault et al. 2017 ⁴	Analytic essay	Global	Qualitative	Data access
Gilbert et al. 2017 ²²	Research article	India	Quantitative	Data collection, data for health service delivery and resource allocation
Githinji et al. 2017 ²³	Research article	Kenya	Quantitative	Data use
Health Metrics Network, World Health Organization 2012. ²⁴	Report	Global	Qualitative	Data access
Higman et al. 2019 ²⁵	Literature review	Global	Qualitative	Data access
Hosseinpoor et al. 2016 ²⁶	Analytic essay	Global	Qualitative	Data collection, data access
Kasambara et al. 2017 ²⁷	Research article	Malawi	Mixed methods	Data collection, data access, data use, data for health service delivery and resource allocation
Krishnan et al. 2010 ²	Research article	India	Qualitative	Data collection, data access, data for health service delivery and resource allocation, relationship between data and accountability of the health system, relationship between data and health outcomes
Llop-Gironés et al. 2019 ²⁸	Research article	Mozambique	Qualitative	Data collection, data for health equity
Madanian et al. 2019 ²⁹	Literature review	India	Qualitative	Data for health equity

(continued)

Table 1. Continued.

Citation	Type of article	Setting	Research method	Topics addressed
Mathers et al. 2009 ³⁰	Analytic essay	Multi-country	Qualitative	Data collection
The Maternal and Child Survival Program 2019. ³¹	Report	Global	Qualitative	Data use
Meachael and Edelman 2019 ³²	Advocacy article	Multi-country	Mixed methods	Data access, data for health service delivery and resource allocation, data for health equity
Mishra et al. 2019 ³³	Report	Global	Mixed methods	Data collection, data access, data for health equity
Moxon et al. 2015 ³⁴	Report	Global	Qualitative	Data collection
Ndabarora et al. 2014 ⁵	Systematic review	Multi-country	Qualitative	Data collection, data access, data for health service delivery and resource allocation, data for health equity
Nguyen et al. 2017 ³⁵	Research article	Vietnam	Quantitative	Data collection, data for health service delivery and resource allocation
Nichols et al. 2019 ³⁶	Systematic review	Multi-country	Mixed methods	Data collection, data access
Nicol et al. 2016 ³⁷	Research article	South Africa	Quantitative	Data collection
Nicol et al. 2017 ³⁸	Research article	South Africa	Qualitative	Data collection, data for health equity
Nutley 2012 ³⁹	Theoretical article	Global	Qualitative	Data for health equity
Open Data Watch 2018. ⁴⁰	Report	Multi-country	Mixed methods	Data access, data use
Open Data Watch 2018 ⁴¹	Analytic essay	Global	Qualitative	Data use
The Partnership in Statistics for Development in the 21st Century (PARIS21) 2019 ⁴²	Report	Global	Mixed methods	Data collection, data access, data use
Oluoch et al. 2015 ⁴³	Research article	Kenya	Quantitative	Data access, data for health service delivery and resource allocation
PATH, World Health Organization, National Institute of Hygiene and Epidemiology 2013 ⁴⁴	Research brief	Vietnam	Quantitative	Data collection, data for health service delivery and resource allocation
PATH, Pan American Health Organization 2019 ⁴⁵	Systematic review	Global	Qualitative	Data collection, data access, data use, data for health service delivery and resource allocation, relationship between data and health outcomes
PATH Tanzania 2018 ⁴⁶	Report	Tanzania	Mixed methods	Data collection, data use, relationship between data and health outcomes

(continued)

Table 1. Continued.

Citation	Type of article	Setting	Research method	Topics addressed
PATH Zambia 2018 ⁴⁷	Report	Zambia	Mixed methods	Data collection, data access, data use, data for health service delivery and resource allocation, relationship between data and monitoring and accountability of the health system
Phillips et al. 2015 ¹	Research article	Global	Quantitative	Data collection, data use, data for health service delivery and resource allocation, data and health equity
Puttkammer et al. 2016 ⁴⁸	Research article	Haiti	Mixed methods	Data collection, data use, data for health service delivery and resource allocation
Qureshi 2016 ⁴⁹	Editorial	Global	Qualitative	Data collection, data access, data for health equity
Roomaney et al. 2017 ⁵⁰	Literature review	South Africa	Qualitative	Data collection, data for health service delivery and resource allocation
Sahay et al. 2018 ⁵¹	Research article	India, Tajikistan	Qualitative	Data collection, data for health service delivery and resource allocation
SDSN TRenDS and Open Data Watch ⁵²	Analytic essay	Global	Qualitative	Data collection
Setel et al. 2007 ⁵³	Research article	Global	Quantitative	Data collection, data access, data use, data for health service delivery and resource allocation, relationship between data and health outcomes
Wang et al. 2012 ⁵⁴	Research article	Global	Quantitative	Data access
Wenz and Abouzahr 2017 ⁵⁵	Report	Global	Qualitative	Data collection, data access, data for health equity
WHO Forum on Health Data Standardization and Interoperability 2013. ⁵⁶	Report	Global	Qualitative	Data access, data use
World Bank, World Health Organization 2014 ⁵⁷	Report	Global	Qualitative	Data collection, data access, data for health equity
Yourkavitch et al. 2016 ⁵⁸	Research article	Malawi	Mixed methods	Data collection, data for health service delivery and resource allocation

Integrity of data recorded. Nearly half of the articles (46%) noted the integrity of data recorded in LMICs influenced the strength of statistical systems, particularly those reporting on vital events. These data issues, including timeliness, completeness, and conformity, impact the reliability of estimates of mortality and morbidity. One systematic review found that missing data within LMIC health information systems ranged from 19% to 50%, and that only 50% of

countries had data that were at least 80% accurate.³ In addition, several studies cited nonconformity—that is, differences in how data on the same outcome are routinely collected and recorded—as a key barrier to aggregating and comparing health outcomes across, and sometimes within, countries.¹⁴ Nine of ten key informants also discussed the manner in which data are recorded as a core factor underpinning discrepancies between data quality in

Table 2. Factors influencing data collection and supporting evidence.

Theme	Supporting articles	Supporting key informants	Main findings
Data collection (33 of 50 articles)			
Integrity of data recorded	23	9	Data accuracy, reliability, and completeness affect LMICs' ability to make well-informed, data-driven decisions regarding health policy and resource allocation.
Mode of data collection and storage	18	2	Different levels of the health system may collect data through different modes, such as claims, self-administered questionnaires, and face-to-face interviews. Each mode introduces different types of biases. Such nonconformity in mode adds complexity when aggregating and interpreting data for decision-making at higher levels of the health system.
Population sampled	13	5	Some populations are disproportionately underrepresented in data across the globe, particularly women, children, and vulnerable groups, limiting the accuracy of health estimates (eg, burden of disease) for these populations and overall in LMICs.
Benefit or burden of collection	4	2	Lack of harmonization in data reporting requirements between multiple vertical programs and funders result in duplicative data collection, particularly among community health workers, and difficulty comparing data across programs.

LMICs and HICs. Several articles concluded that to participate in data-driven decision-making and hold governing bodies accountable for providing social services, LMICs need data systems that can adequately link health outcomes and quality of care.^{1,14,17,28,30,55}

Mode of data collection and storage. Eighteen articles highlighted how the mode of data collection contributes to broader issues of data validity, or whether the way data are collected captures the true measure of interest. Two key informants and three articles described how the same data element (for example, information on referrals or drug stockouts) collected by staff in different positions and differentially stored (electronic and paper) across the health system led to decision-making based on inaccurate or incomplete data, thus tenuously grounding the decisions in data at all.^{3,21,58} Lack of standardization around the mode of data collection often makes organizing, integrating, and retrieving voluminous medical and patient information difficult, particularly in paper-based systems.^{3,15} Computerized and mobile data collection systems help ensure more accurate and timely data due to inbuilt systems that cross-check completeness; these systems also considerably reduce time spent on data entry.^{2,13,45} In addition, studies suggest that using mobile phones for data entry saves considerable travel time for community health workers who would otherwise need to transfer paper-based client information to electronic systems at health facilities,

allowing them to spend more time on direct health service delivery.¹³

Population sampled. Thirteen articles and five key informants specifically discussed the underrepresentation of women, children, marginalized groups, vulnerable, and hard-to-reach populations, such as refugees, in CRVS systems, District Health Information Software (DHIS), censuses, and other household surveys.^{7,52,55} These data gaps cause inaccuracy in population-level health estimates (eg, morbidity, mortality, cause of death) and render analyses such as assessing equity in outcomes difficult or impossible to reliably conduct because large sample sizes of smaller populations are needed to generate sufficient statistical power.³⁴ One study in Mozambique found that none of the country's 27 published indicators related to Sustainable Development Goal 3 had enough disaggregated information to conduct a full analysis along the dimensions of sex, age, socioeconomic position, geographic area, and race or ethnicity.²⁸ Two studies described how published data are insufficiently robust for conducting health equity analyses across geographies and health topics.^{6,26}

Benefit and burden of data collection. Four articles addressed how health workers charged with data collection require additional training and support to develop core competencies to recognize, anticipate, and correct for data issues during the collection process.^{38,50} However, these articles

Table 3. Key themes on data access and supporting evidence.

Theme	Supporting articles	Supporting key informants	Main findings
Data access (27 of 50 articles)			
Interoperability	18	6	Most data systems in LMICs operate in silos and without mechanisms for communication to and across a variety of ministries, funders, and vertical programs. No one party can view or leverage collected data in their entirety for societal good.
Interface	7	5	Data in LMICs are frequently reported in text or PDF format, rather than more accessible data types, such as comma-separated value (CSV) files. LMIC stakeholders also have limited access to geocoded data, hindering geospatial displays with which users would find easier to interact. In cases where data are available electronically and allow for more sophisticated analysis, potential users at the frontline level often do not have the technology or data know-how to access the data.
Policies governing data sharing	9	4	In contrast to HICs, LMICs do not have legal frameworks around data sharing and interoperability standards, such as recent policy governing ethics and data privacy standards. The absence of these policies has, in some instances, led to concerns around patient confidentiality and data sharing, though several LMICs have made improvements in adherence to data dissemination standards over the past 15 years.

highlighted that most health workers have little supervision or feedback for data collection and often have poor attitudes toward data tasks, with few incentives to recognize the benefits of the data to them or to engage in and manage the burden of collection. Two key informants described how data collectors (eg, community health workers, health facility personnel, program coordinators) may be required to use different tools to enter similar—yet slightly different—data elements, sometimes across multiple vertical programs, resulting in the loss of morale and effort due to duplicative and time-consuming data entry. Data collectors in remote areas are also known to experience resource limitations and logistical difficulties when sending data to higher levels of the health system, leading to delayed submission of reports and misplaced or fabricated data.²⁷ Burdensome data entry requirements can also lead community health workers, clinicians, and nurses to juggle data collection with their patient care duties, often at the expense of providing quality health services.^{15,27}

Access to data for health and health equity

Approximately half (27) of the relevant articles touched on how data access affects health and health equity in LMICs. Dimensions of data access examined included interoperability of systems to enable information sharing, interface through which data become accessible, and policies governing data sharing (Table 3). Two of the three dimensions

were related to the design of the system and the data format and systems' ability to "talk" to one another. The third dimension related to policies governing data sharing determines various groups' ability to sharing collected data and gain analytic insights from them.

Interoperability. Eighteen articles and six key informants described how integrated systems could expand data access to more users and increase coordination between sectors to gain action-oriented insights for health and health equity. The key informants highlighted that current environments silo data within programs and limit their use to only groups with permission. The articles primarily called for more interoperable data systems that allow for exchange of back-end data within and across organizations to reduce duplication and increase sharing and accessibility. Developing them requires a deliberate system architecture with input across many domains of the health system (eg, health staff, commodities, logistics, progress tracking, financing).⁴⁹

Interface. Seven articles and two key informants addressed the front-end interface through which users access data, highlighting the need for more spatial displays of geocoded information.⁷ Key informants described how data in LMICs are often published in non-interactive text reports or PDF form, further limiting their accessibility and use within the health system, as well as to the

Table 4. Key themes on data use and supporting evidence.

Theme	Supporting articles	Supporting key informants	Main findings
Data use (19 of 50 articles)			
Demand for data across the health system	13	5	Frontline workers at lower levels of the health system overwhelmingly function as producers of data for monitoring purposes, but less frequently view themselves as consumers of data. As a result, they do not seek to use health data for actionable insights to improve health service delivery and resource allocation for improved health outcomes and greater health equity.
Technical capacity	10	7	LMICs typically lack robust training on data use to build the necessary underlying skills to conduct sophisticated data analysis. Moreover, individuals who do possess strong data analysis skills do not receive sufficient incentive to work in the public sector.

public. As more LMICs have introduced computerized systems (eg, EIR, LMIS, DHIS2, HMIS, and/or CDSS), one study found that such systems increase health workers' ability to interpret routine data, particularly at the district and province levels of the health system.⁴⁵ At the same time, having an electronic interface does not guarantee access for all—three studies noted that many LMIC health workers cannot log into computers or the Internet at work, impacting not only their ability to send and transmit data to higher levels of the health system for health decision-making but also to use processed data for their own work.^{21,33,55} However, one study described how mobile devices serve as a mechanism to overcome technology constraints among frontline workers and improve data sharing and access across levels of the health system.³³

Policies governing data sharing. One expert noted that formal guidance and policies can encourage individuals, organizations, and agencies to share data. A recent analysis showed a 72-percentage point increase in the number of countries adhering to International Monetary Fund data dissemination standards over the past 15 years, including some LMICs.⁴²

To date, however, LMICs have generally lacked legal frameworks around data ethics, sharing, and interoperability standards, such as modern policies governing ethics and data privacy standards. Citing a lack of culture around data sharing and broad skepticism around data privacy and security, three articles and one expert described how countries with existing governmental and organizational data policies tend to focus on *limiting* data flows across countries, sectors, and entities.^{5,33,55}

Data use for health and health equity

Nineteen articles addressed inequities in data use. Components of data use included demand for data across the health system and technical capacity for data use (Table 4). Both of these components require knowledge of potential data users, including awareness of the data, understanding of their possible uses, and technical expertise to generate insights.

Demand for data across the health system. Four articles demonstrated how perceptions of improved data quality increased demand for and interest in data as an analytic tool in health decision-making.^{13,45–47} However, 13 articles and five key informants highlighted the historically low demand for data in LMICs. Although many frontline workers in LMICs produce data to meet reporting requirements, few simultaneously function as data consumers. Articles described how several ongoing factors inhibited demand for data—perceptions of poor data quality, lack of motivation, burden of workload outside of data duties, and limited incentive structures for sophisticated data use—particularly at the frontlines of LMIC health systems.^{21,23,45}

Technical capacity. Ten of the 19 articles related to data use and seven key informants cited technical capacity as a factor leading to effective data use. Key informants expressed broad agreement with the importance of improved analytic capacity across all levels of the health system, noting that current training on data use—when it exists—is infrequent and is typically limited to generating simple analytics. Common challenges related to technical capacity included knowledge gaps, lack of training and

refresher courses, lack of data quality reviews, and improper comprehension of surveillance terms.^{27,57} A recent assessment of statistical capacity performance across world regions showed that Europe and the Americas lead, with Africa and Oceania (excluding Australia and New Zealand) trailing far behind.⁴² Most countries, however, have improved their use of statistics in national planning and policymaking since 2000, and two evaluations assessing interventions to improve technical capacity for immunization data quality found that training efforts and systems modifications improved the timeliness and completeness of, as well as overall confidence in, the data, with early indications that these data quality improvements paved the way for improved data use.^{46,47}

Data and improved health service delivery, outcomes, and equity

Our scoping review uncovered 19 articles associating health data with improved health service delivery. Eleven articles connected health data to improvements in health center operations, specifically linking greater accessibility and use of data to improvements in maintaining drug stocks, administering timely vaccinations, and aligning budgets to health service needs.^{22,35,53} Four articles linked health data with improvements in provider performance and facility operations, describing how computerized HMIS systems and mobile reminders improved drug stock management and resulted in fewer overdue client visits. Use of mobile decision support tools also showed improved adherence to guidelines and protocols, resulting in improved vaccination coverage rates and quality of care.^{2,5,13,47}

Although key informants tended to believe in a virtuous cycle between data and improved health outcomes, they overwhelmingly agreed little evidence demonstrated the link between the two. Among the articles attempting to link data to health improvements, none were able to demonstrate a link between data and health outcomes, and two postulated that data improvements could lead to health equity with little to no evidence to support that connection.

Discussion

Data equity requires all stakeholders to have high quality data available for use, the capacity to use it effectively, and to regularly use it to achieve results. Some articles offered frameworks to improve the quality and quantity of data available for health-related decision-making in LMICs. Although the majority of these frameworks addressed specific mechanisms through which data could be improved, very few expanded the framework to the point of showing how data improvements could lead to changes in health service delivery or outcomes. Our review found little evidence supporting the connection

between data and health or equity outcomes, yet it did emphasize where changes in data collection, access, and use might lead to better outcomes in LMICs.

To close the data gap between LMICs and HICs, our findings emphasized more inclusive methods of data collection, systems allowing for open data sharing, and sustained human capital as key pathways to move from an inequitable to equitable data landscape. For data collection, LMICs have an opportunity to expand the subset and quality of data they currently collect for health-related decision-making. To increase the quality of data collected, LMICs can select streamlined modes of collection and enhance trainings for people charged with data collection. Transitioning from a paper-based to electronic data storage system has increased data access across geographies for some, but making data fully accessible to those who need them will require policies to regulate data sharing and infrastructure to improve interoperability across ministries, agencies, and organizations. Key informants point to decreasing time-lags in data dissemination and embedding analytic capacity in data sharing modes as key to increasing relevance and usage of data.

Such changes to increase data equity must be accompanied by knowledge and skills to fully seek and use these data. LMICs can conduct an assessment of their data capacity and quality to identify strengths and inequities in their data collection, access, and use. Then, they can identify the partnerships and skills needed to leverage the strengths and fill in gaps to catalyze change and forward representative and useful data for all. Without necessary knowledge and skills, a state of data inequity will only perpetuate in LMICs.

An implied consequence of relying on largely flawed information is subjecting decisions on important matters, such as health resource allocation, to considerable error and potentially worsening health inequities among the populations in question. Given this and the commitment global health stakeholders have placed on data and data improvement for health in recent history, it is worthy to continue examining whether efforts to close the gap in data between LMICs and HICs yield improvements in health outcomes and lay the foundation for health equity.

To guide this examination, we offer a conceptual framework of data equity for health and health equity (DE4HHE) based on our scoping review. The DE4HHE framework outlines the uncovered disparities in LMIC data, demonstrates how various aspects of these disparities relate to one another, and—once collectively improved—will bridge the gap to create a state of data equity (Figure 2). It offers global health stakeholders a model along which to measure progress in the pathway to data equity and to assess the true relationship between data equity and health equity.

Our search strategy deliberately excluded certain terms to manage the scope of the review, which may have impacted the comprehensiveness of our results. Notably,

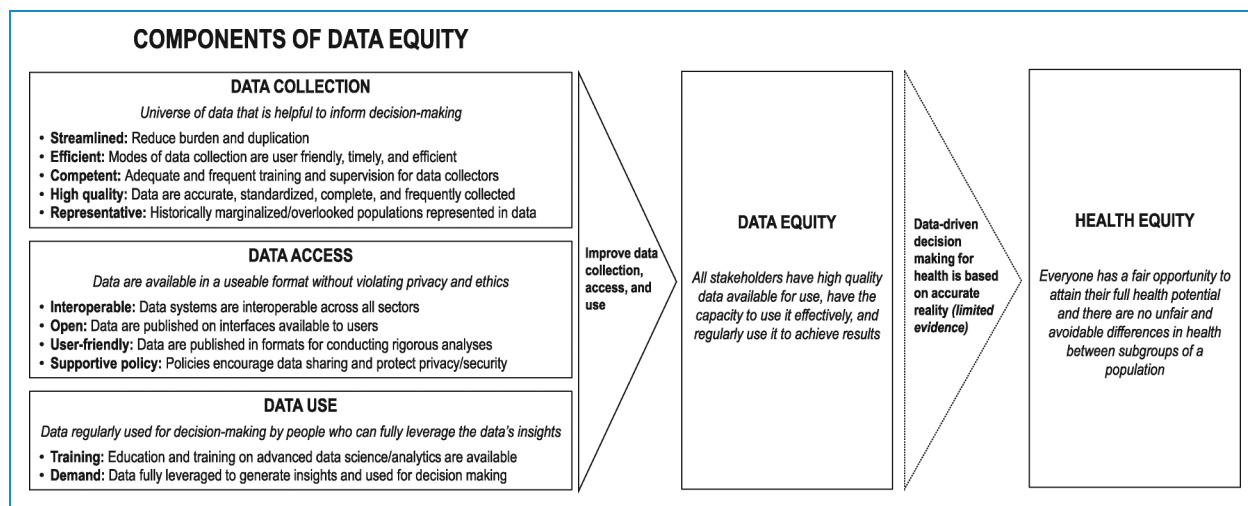


Figure 2. Conceptual framework of components of data equity for health and health equity.

we elected to exclude noncommunicable diseases (NCDs) and surveillance among our search terms. However, some of the articles included in our review profiled interventions to improve data collection, access, and use related to NCDs and surveillance systems in LMICs. We excluded these specific search terms out of concern that the included research would encompass an onerous number of studies with only tangential relevance to data equity. In addition, the study was funded and conducted in 2019, with a look back period of 10 years. As a result, the literature does not capture more recent data best practices and innovations that have been surfaced or accelerated due to the current data context brought about by COVID-19.

In addition to potentially excluding articles that may have had relevance to our review, very few of the articles that we did review were assessed to be of high evidence. A number of the topics we explored over the course of our review, such as policies governing data access, are not easily translated into rigorous studies to generate high quality evidence. Therefore, the conclusions we were able to draw were overwhelmingly based on articles assessed to have low evidence underpinning their claims. Although a clear portrait regarding health data inequities nonetheless emerged over the course of our review and interviews with key informants, we recognize that rooting our understanding of the landscape in low quality evidence may have impacted the validity of our assumptions.

Our findings were also shaped by our selection of key informants, all but one of whom hail from the United States. While they have all contributed valuable expertise to the field and have conducted extensive research in LMICs, their perspectives are nevertheless shaped by their location and affiliation with institutions in the Global North. Future work in this domain could be strengthened by incorporating a wider range of insights from key informants who have spent their lives living and working in

LMICs, such as health information system data managers and data collectors and users at all levels of the health system.

Finally, given the level of evidence available during the study period and the wide variety of geographies and types of data interventions reviewed, our findings were ultimately quite broad. This scope has facilitated our ability to create a conceptual framework with potential applicability to a wide variety of contexts. We acknowledge, however, that our conclusions inevitably will not represent the data landscape in specific geographies within LMICs and should only be used as a broad depiction of common health data inequities across these contexts.

Conclusion

With the importance of data-driven decision-making increasingly emphasized across all sectors, the public health and health care fields have seen the emergence of big data and predictive analytic tools that take data and transform them into insights for action.⁵⁹ Our scoping review revealed select cases where data improved health system management and outcomes, such as timely vaccine administration, greater adherence to guidelines, and more active drug stock management. However, the evidence on data's direct impact on health service delivery, behaviors, and outcomes is still emerging. The DE4HHE conceptual framework identified the core aspects of data inequity in LMICs to be further examined to understand which combination of these aspects, if improved, might catalyze better health for *all*. The next steps include identifying measures to assess each of these data equity conceptual framework components and assessing the association of these components to positive health outcomes, especially in populations experiencing high burden of risk and disease.

Acknowledgements: Many thanks to colleagues Divya Vohra and Joshua Claxton for assisting the authors with article screening.

Contributorship: SO conceived the study. SO, ST, and AS designed the search strategy. All authors screened the full text of potentially relevant articles. SO and ST wrote the first draft of the manuscript. All authors critically revised the manuscript and contributed to subsequent iterations.

Declaration of conflicting interests: The authors declare that there is no conflict of interest. Our research was funded by the Rockefeller Foundation (grant number 9661). The funder of the study reviewed and provided feedback on the key informant interview questions and the high-level approach of the study. However, the authors made the final decisions on all aspects of the research, had full access to all the data in the study, and had final responsibility for the final content decision to submit for publication.

Funding: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Rockefeller Foundation (grant number 9661).

Guarantor: SO

ORCID iDs: Sasigant O'Neil  <https://orcid.org/0000-0003-1780-4016>
 Sydney Taylor  <https://orcid.org/0000-0001-5495-6866>
 Anitha Sivasankaran  <https://orcid.org/0000-0003-1162-4268>

References

- Phillips DE, AbouZahr C, Lopez AD, et al. Are well functioning civil registration and vital statistics systems associated with better health outcomes? *Lancet* 2015; 386: 1386–1394.
- Krishnan A, Nongkynrih B, Yadav K, et al. Evaluation of computerized health management information system for primary health care in rural India. *BMC Health Serv Res* 10, 310 (2010). <https://doi.org/10.1186/1472-6963-10-310>
- Fenenga C and de Jager A. Cordaid-IICD health programme Uganda: health management information systems as a tool for organisational development. *The Electronic Journal of Information Systems in Developing Countries* 2007; 31: 1–14.
- Flahault A, Geissbuhler A, Guessous I, Guérin PJ, Bolon I, Salathé M and Escher G. Precision global health in the digital age. *Swiss Med Wkly.* 2017 Apr 8;147:w14423. doi: 10.4414/SMW.2017.14423. PMID: 28421566.
- Ndabarora E, Chippis JA and Uys L. Systematic review of health data quality management and best practices at community and district levels in LMIC. *Information Development* 2014; 30: 103–120.
- Clouse K, Phillips T and Myer L. Understanding data sources to measure patient retention in HIV care in sub-saharan Africa. *Int Health* 2017; 9: 203–205.
- Kruk ME, Gage AD, Arsenaault C, et al. High-quality health systems in the sustainable development goals era: time for a revolution. *Lancet Glob Health* 2018; 6: e1196–e1252.
- Chowdhury R, Kumari L and Panda S. The role of health information system in maternal and child health-care services. *Asian Journal of Pharmaceutical and Clinical Research* 2017; 10: 29–31.
- Kerber KJ, Mathai M, Lewis G, et al. Counting every stillbirth and neonatal death through mortality audit to improve quality of care for every pregnant woman and her baby. *BMC Pregnancy Childbirth* 15, S9 (2015). <https://doi.org/10.1186/1471-2393-15-S2-S9>.
- Fenech M, Strukelj N and Buston O. *Ethical, social, and political challenges of artificial intelligence in health* [Internet]. April 2018. Available from: <https://wellcome.org/sites/default/files/ai-in-health-ethical-social-political-challenges.pdf>.
- World Health Organization. Regional Office for the Western Pacific. *Improving data quality: A guide for developing countries* [Internet]. 2003. Available from: <https://apps.who.int/iris/handle/10665/206974>.
- Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018; 169: 467–473.
- Agarwal S, Perry HB, Long L-A, et al. Evidence on feasibility and effective use of mHealth strategies by frontline health workers in developing countries: systematic review. *Trop Med Int Health* 2015; 20: 1003–1014.
- Akachi Y and Kruk ME. Quality of care: measuring a neglected driver of improved health. *Bull World Health Organ* 2017; 95: 465–472.
- Bhavaraju SR. From subconscious to conscious to artificial intelligence: a focus on electronic health records. *Neuro India* 2018; 66: 1270–1275.
- Burke J. *Health analytics: gaining the insights to transform health care*. Hoboken, NJ: John Wiley & Sons, 2013.
- Buvinic M, Furst-Nichols R and Koolwal G. Mapping gender data gaps [Internet]. March 2014. Available from: https://data2x.org/wp-content/uploads/2019/05/Data2X_MappingGenderDataGaps_FullReport.pdf.
- Carney TJ and Weber DJ. Public health intelligence: learning from the ebola crisis. *Am J Public Health* 2015; 105: 1740–1744.
- Chandy L and Zhang C. *Stuffing data gaps with dollars: what will it cost to close the data deficit in poor countries?* [Internet]. August 2015. Available from: <https://www.brookings.edu/opinions/stuffing-data-gaps-with-dollars-what-will-it-cost-to-close-the-data-deficit-in-poor-countries/>.
- Devex P. *Digital health solutions: lessons and best practices from implementers* [Internet]. 2019. Available from: https://pages.devex.com/rs/685-KBL-765/images/REPORT_Digital_Health_Solutions.pdf.
- Endriyas M, Alano A, Mekonnen E, et al. Understanding performance data: health management information system data accuracy in southern nations nationalities and People's Region, Ethiopia. *BMC Health Serv Res* 2019; 19: 175.
- Gilbert S. *Navigating digital transformation: 8 trends defining the future of global digital health* [Internet]. December 2017. Available from: <https://digitalsquare.org/blog/2017/12/6/navigating-digital-transformation8-trends-defining-the-future-of-global-digital-health>.
- Githinji S, Oyando R, Malinga J, et al. Completeness of malaria indicator data reporting via the District Health Information Software 2 in Kenya, 2011–2015. *Malar J* 16, 344 (2017). <https://doi.org/10.1186/s12936-017-1973-y>

24. Health Metrics Network, World Health Organization. Framework and standards for country health information systems [Internet]. 2012(Second edition). Available from: https://www.who.int/healthinfo/country_monitoring_evaluation/who-hmn-framework-standards-chi.pdf.
25. Higman S, Dwivedi V, Nsagurwe A, et al. Designing interoperable health information systems using enterprise architecture approach in resource limited countries: a literature review. *Int J Health Plann Manage* 2019; 34: e85–e99.
26. Hosseinpoor AR, Bergen N, Schlottheuber A, et al. Data resource profile: WHO health equity monitor (HEM). *Int J Epidemiol* 2016; 45: 1404e–1405e.
27. Kasambara A, Kumwenda S, Kalulu K, et al. Assessment of implementation of the health management information system at the district level in southern Malawi. *Malawi Med J* 2017; 29: 240–246.
28. Llop-Gironés A, Cash-Gibson L, Chicumbe S, et al. Health equity monitoring is essential in public health: lessons from Mozambique. *Global Health* 15, 67 (2019). <https://doi.org/10.1186/s12992-019-0508-4>
29. Madanian S, Parry DT, Airehrour D and Cherrington M. mHealth and big-data integration: promises for healthcare system in India. *BMJ Health Care Inform* 2019 Sep; 26(1): e100071. doi: 10.1136/bmjhci-2019-100071. PMID: 31488497; PMCID: PMC7062344.
30. Mathers CD, Boerma T and Ma Fat D. Global and regional causes of death. *Br Med Bull* 2009; 92: 7–32.
31. The Maternal and Child Survival Program. Fostering use of routine reproductive, maternal, newborn, and child health data at the point of care: experience applying data dashboards in low- and middle-income countries [Internet]. 2019. Available from: <https://www.mcsprogram.org/resource/fostering-use-of-routine-reproductive-maternal-newborn-and-child-health-data-at-the-point-of-care/>.
32. Michael P and Edelman JK. Global digital health index: the state of digital health 2019 [Internet]. April 2019. Available from: <https://static1.squarespace.com/static/5ace2d0c5cfd792078a05e5f/t/5d4dcb80a9b3640001183a34/1565379490219/State+of+Digital+Health+2019.pdf>.
33. Mishra SR, Lygidakis C, Neupane D, et al. Combating non-communicable diseases: potentials and challenges for community health workers in a digital age, a narrative review of the literature. *Health Policy Plan* 2019; 34: 55–66.
34. Moxon SG, Ruysen H, Kerber KJ, et al. Count every newborn; a measurement improvement roadmap for coverage data. *BMC Pregnancy Childbirth* 15, S8 (2015). <https://doi.org/10.1186/1471-2393-15-S2-S8>
35. Nguyen NT, Vu HM, Dao SD, et al. Digital immunization registry: evidence for the impact of mHealth on enhancing the immunization system and improving immunization coverage for children under one year old in Vietnam. *Mhealth* 2017 Jul 19; 3: 26. doi: 10.21037/mhealth.2017.06.03. PMID: 28828373; PMCID: PMC5547172.
36. Nichols EK, Raganathan NW, Raganathan B, et al. A systematic review of vital events tracking by community health agents. *Glob Health Action* 2019; 12(1), 1597452. <https://doi.org/10.1080/16549716.2019.1597452>
37. Nicol E, Dudley L and Bradshaw D. Assessing the quality of routine data for the prevention of mother-to-child transmission of HIV: an analytical observational study in two health districts with high HIV prevalence in South Africa. *Int J Med Inf* 2016; 95: 60–70.
38. Nicol E, Bradshaw D, Uwimana-Nicol J, et al. Perceptions about data-informed decisions: an assessment of information-use in high HIV-prevalence settings in South Africa. *BMC Health Serv Res* 17, 765 (2017). <https://doi.org/10.1186/s12913-017-2641-1>
39. Nutley T. Improving data use in decision making: an intervention to strengthen health systems [Internet]. August 2012. Available from: https://www.measureevaluation.org/resources/publications/sr-12-73/at_download/document.
40. Open Data Watch. Measuring data use: an analysis of data portal web traffic [Internet]. 2018. Available from: <https://opendatawatch.com/wp-content/uploads/2018/11/Measuring-Data-Use-Report.pdf>.
41. Open Data Watch. The data value chain: moving from production to impact [Internet]. 2018. Available from: https://opendatawatch.com/wp-content/uploads/2018/03/Data_Value_Chain-WR-1803126.pdf.
42. The Partnership in Statistics for Development in the 21st Century (PARIS21). Statistical Capacity Development Outlook 2019 [Internet]. 2019. Available from: <https://paris21.org/sites/default/files/inline-files/Statistical%20Capacity%20Development%20Outlook%202019.pdf>.
43. Oluoch T, Muturi D, Kiriinya R, et al. Do interoperable national information systems enhance availability of data to assess the effect of scale-up of HIV services on health workforce deployment in resource-limited countries? *Stud Health Technol Inform* 2015; 216: 677–681.
44. PATH, World Health Organization, National Institute of Hygiene and Epidemiology. Vietnam’s immunization registries go online [Internet]. Available from: https://path.azureedge.net/media/documents/TS_opt_viet_imm_reg_br.pdf. 2013.
45. PATH, Pan American Health Organization. A realist review of what works to improve data use for immunization: evidence from low- and middle-income countries [Internet]. 2019. Available from: https://reliefweb.int/sites/reliefweb.int/files/resources/IDEA_WhitePaper_English.pdf.
46. PATH Tanzania. BID Initiative final evaluation report: immunization data quality and use in Tanzania [Internet]. June 2018. Available from: http://bidinitiative.org/wp-content/uploads/BID-Initiative_ME-Endline-Report_Tanzania_FINAL_20August2018.pdf.
47. PATH Zambia. BID Initiative Zambia midline report [Internet]. June 2018. Available from: https://bidinitiative.org/wp-content/uploads/BID-Initiative_ME-Midline-Report_Zambia_FINAL.pdf.
48. Puttkammer N, Baseman JG, Devine EB, et al. An assessment of data quality in a multi-site electronic medical record system in Haiti. *Int J Med Inf* 2016; 86: 104–116.
49. Qureshi S. Creating a better world with information and communication technologies: health equity. *Information Technology for Development* 2016; 22: 1–14.
50. Roomaney RA, Pillay-van Wyk V, Awotiwon OF, et al. Availability and quality of routine morbidity data: review of studies in South Africa. *J Am Med Inform Assoc* 2017; 24: e194–e206.
51. Sahay S, Nielsen P and Latifov M. Grand challenges of public health: how can health information systems support facing them? *Health Policy Technol* 2018; 7: 81–87.

52. SDSN TReNDS and Open Data Watch. Civil registration and vital statistics benefit health, child protection, and governance [Internet]. 2018. Available from: <https://www.data4sdgs.org/sites/default/files/2018-12/CRVS%20Case%20Study.pdf>.
53. Setel PW, Macfarlane SB, Szreter S, et al. A scandal of invisibility: making everyone count by counting everyone. *Lancet* 2007; 370: 1569–1577.
54. Wang H, Dwyer-Lindgren L, Lofgren KT, et al. Age-specific and sex-specific mortality in 187 countries, 1970-2010: a systematic analysis for the global burden of disease study 2010. *Lancet* 2012; 380: 2071–2094.
55. Wenz K and Abouzahr C. Better data for women and children: strengthening civil registration and vital statistics across the continuum of care. Technical meeting report [Internet]. 2017. Available from: https://www.healthdatacollaborative.org/fileadmin/uploads/hdc/Documents/Working_Groups/Better_Data_for_Women_and_Children_Strengthening_CRVS.pdf.
56. World Health Organization. WHO Forum on Health Data Standardization and Interoperability [Internet]. 2013. Available from: https://www.who.int/ehealth/WHO_Forum_on_HDSI_Report.pdf?ua=1.
57. World Bank, World Health Organization. Global Civil Registration and Vital Statistics Scaling up Investment Plan 2015–2024 [Internet]. 2014. Available from: <https://openknowledge.worldbank.org/handle/10986/18962>.
58. Yourkavitch J, Zalisk K, Prosnitz D, et al. How do we know? An assessment of integrated community case management data quality in four districts of Malawi. *Health Policy Plann* 2016; 31: 1162–1171.
59. Fruchterman J. Using data for action and for impact. *Stanford Social Innovation Review* 2016; 20, Available from: https://ssir.org/articles/entry/using_data_for_action_and_for_impact.

Appendix 1

Websites included in targeted grey literature search and search terms used

Organization	Web address
Bloomberg Data for Health Initiative	https://www.bloomberg.org/program/public-health/data-health/
Bloomberg Philanthropies Data Impact Program	http://www.d4hdataimpact.org/
Better Immunization Data (BID) Initiative	http://bidinitiative.org/
The Brookings Institution	https://www.brookings.edu/
Data2x	https://data2x.org/
	http://opendataenterprise.org/

(continued)

Continued.

Organization	Web address
Center for Open Data Enterprise	
Community Health Roadmap	https://www.communityhealthroadmap.org/
Development Gateway	https://www.developmentgateway.org/
Global Partnership for Sustainable Development Data	http://www.data4sdgs.org/
Global Digital Health Index	https://www.digitalhealthindex.org/
Harvard Data Science Review	https://review.datascience.harvard.edu/
Institute for Health Metrics and Evaluation	http://www.healthdata.org/
International Development Research Centre	https://www.idrc.ca/en
Jhpiego	https://www.jhpiego.org/
MEASURE Evaluation	https://www.measureevaluation.org/
Open Data Watch	https://opendatawatch.com/
Stanford Social Innovation Review	https://ssir.org/
Sustainable Development Goals Knowledge Platform	https://sustainabledevelopment.un.org/sdgs
Thematic Research Network on Data and Statistics (TReNDS)	https://www.sdsntrends.org/
Understanding Patient Data	https://understandingpatientdata.org.uk/case-studies
World Federation of Public Health Associations	https://www.wfpha.org/
The World Bank	https://www.worldbank.org/
The World Economic Forum	https://www.weforum.org/
World Health Organization	https://www.who.int/

Search terms:

("data gaps" OR "data representativeness" OR "data openness" OR "data completeness" OR "data availability" OR "civil registration" OR "data accuracy") AND ("global health" OR "health equity" OR "mHealth") AND ("low-income countries" OR "middle-income countries" OR "developing countries") AND ("immunization" OR "vaccination" OR HIV OR AIDS OR malaria)

("data infrastructure" OR "data analytics" OR "digital disruption" OR "big data" OR "data policy" OR "data ecosystems" OR "data systems") AND ("global health" OR "health equity" OR "mHealth") AND ("low-income countries" OR "middle-income countries" OR "developing countries") AND ("immunization" OR "vaccination" OR HIV OR AIDS OR malaria)

Appendix 2**Search strategy for peer reviewed literature**

("Data equity" OR "digital divide" OR "data gaps" OR "data representativeness" OR "data openness" OR "data completeness" OR "data availability" OR "civil registration and vital statistics" OR "data infrastructure" OR "data analytics" OR "data science" OR "digital disruption" OR "internet of things" OR "big data" OR "data policy" OR "data ecosystems" OR "data systems")

AND

("health equity" OR "global health" OR eHealth OR mHealth OR eMedicine OR "health information systems" OR "DHIS2" OR "Electronic Medical Record*" OR "electronic health record*")

AND

("low-income econom*" OR "lower-middle-income econom*" OR "upper-middle-income econom*" OR "transforming econom*" OR "non-transforming econom*" OR "developing econom*" OR "second world econom*" OR "third world econom*" OR "poor econom*" OR "emerging econom*" OR "low-income countr*" OR "lower-middle-income countr*" OR "upper-middle-income countr*" OR "transforming countr*" OR "non-transforming countr*" OR "developing countr*" OR "second world countr*" OR "third world countr*" OR "poor countr*" OR "emerging countr*" OR "South Asia" OR "Central Asia" OR "Middle East" OR "North Africa" OR "East Asia" OR "Asia Pacific" OR "Sub-Saharan Africa" OR "Global South" OR "Afghanistan" OR "Albania" OR "Algeria" OR "Angola" OR "Armenia" OR "Azerbaijan" OR "Bangladesh" OR "Belarus" OR "Benin" OR "Bhutan" OR "Bosnia and Herzegovina" OR "Botswana" OR "Bulgaria" OR "Burkina Faso" OR "Burundi" OR "Cape Verde" OR "Cabo Verde" OR "Cambodia" OR "Cameroon" OR "Central African Republic" OR "Chad" OR "China" OR "Comoros" OR "Democratic Congo Republic" OR "Côte

d'Ivoire" OR "Djibouti" OR "Egypt" OR "Equatorial Guinea" OR "Eritrea" OR "Ethiopia" OR "Fiji" OR "Gabon" OR "Gambia" OR "Georgia" OR "Ghana" OR "Guinea" OR "Guinea-Bissau" OR "India" OR "Indonesia" OR "Iran" OR "Iraq" OR "Jordan" OR "Kazakhstan" OR "Kenya" OR "Kiribati" OR "North Korea" OR "Kosovo" OR "Kyrgyzstan" OR "Laos" OR "Lebanon" OR "Lesotho" OR "Liberia" OR "Libya" OR "Macedonia" OR "Madagascar" OR "Malawi" OR "Malaysia" OR "Maldives" OR "Mali" OR "Marshall Islands" OR "Mauritania" OR "Mauritius" OR "Micronesia" OR "Moldova" OR "Mongolia" OR "Montenegro" OR "Morocco" OR "Mozambique" OR "Myanmar" OR "Namibia" OR "Nauru" OR "Nepal" OR "Niger" OR "Nigeria" OR "Pakistan" OR "Papua New Guinea" OR "Philippines" OR "Romania" OR "Russia" OR "Rwanda" OR "Samoa" OR "Sao Tome and Principe" OR "Senegal" OR "Serbia" OR "Sierra Leone" OR "Solomon Islands" OR "Somalia" OR "South Africa" OR "South Sudan" OR "Sri Lanka" OR "Sudan" OR "Swaziland" OR "Syria" OR "Tajikistan" OR "Tanzania" OR "Thailand" OR "Timor-Leste" OR "Togo" OR "Tonga" OR "Tunisia" OR "Turkey" OR "Turkmenistan" OR "Tuvalu" OR "Uganda" OR "Ukraine" OR "Uzbekistan" OR "Vanuatu" OR "Vietnam" OR "West Bank and Gaza" OR "Yemen" OR "Zambia" OR "Zimbabwe" OR "Northern Triangle" OR "Latin America" OR "Caribbean" OR "Latin America and the Caribbean" OR "Central America" OR "El Salvador" OR Guatemala OR Honduras OR Nicaragua OR Bolivia OR Brazil OR Cuba OR "Dominican Republic" OR Colombia OR Ecuador OR Haiti OR Jamaica OR Mexico OR Panama OR Paraguay OR Peru OR Venezuela OR "Antigua and Barbuda" OR Argentina OR Aruba OR Bahamas OR Barbados OR Belize OR "British Virgin Islands" OR "Cayman Islands" OR Chile OR "Costa Rica" OR "Curaçao" OR "Dominica" OR Grenada OR Guyana OR "Puerto Rico" OR "Sint Maarten" OR "St Kitts and Nevis" OR "St Lucia" OR "St Martin" OR "St Vincent and the Grenadines" OR Suriname OR "Trinidad and Tobago" OR "Turks and Caicos Islands" OR Uruguay OR "Virgin Islands")

AND

("health outcomes" OR "community health" OR "data-driven decision-making" OR "information literacy" OR "decision-making" OR policymaking OR "policy making" OR "frontline health" OR "primary health care" OR "universal health coverage" OR "health systems strengthening" OR "maternal and child health" OR immunization OR vaccination OR "infant mortality" OR "maternal mortality" OR "skilled birth attendance" OR "precision health" OR "precision medicine" OR "precision public health")

Appendix 3

Screening Criteria

1. Relevant and useful to the development of the evidence review; that is, has one or more of the following features (and ideally at least two to be screened in):
 - Focused on health data, health information systems, data representativeness, data completeness, data gaps, open data, data access, and/or data equity
 - Focused on health outcomes and/or health equity
 - Focused on community health
 - Insights into the relationship between data use and improved health outcomes and/or improved decision-making
2. Low- and middle-income country context unless the source is particularly strong in other areas.
3. Peer reviewed journal or carefully reviewed report
4. A review of the abstract does not lead to major reservations about the document's use of scientifically-based research standards – that is, “standards that apply rigorous, systematic, and objective methodology to obtain reliable and valid knowledge and present findings and make claims that are appropriate to and supported by the methods employed” (ESRA, Title I of P.L. 107–279). Though opinion pieces may be included to illustrate various points of view in the field.

Appendix 4

Abstraction template fields

Citation; Authors, affiliations, and credentials; Funder/sponsor; Type of article; Peer-reviewed; Sample description and size; Setting; Type of study design; Brief design description; Key data sources; Primary analysis methods; Benefits for development of scoping review; Inequities in data collection; Inequities in data access; Inequities in data use; Data for health service delivery and resource allocation; Relationship between data and monitoring and accountability of the health system; Relationship between data and health outcomes; Key findings relevant to opportunities to achieve health equity; Study limitations and cautions; Assessment of objectivity; Initial rating of research quality; Notes; Indicators noted in the article; Interventions cited in the article; Agencies, organizations, and/or resources to note; Citations from article; Gaps.

Appendix 5

Statements used for resonance testing with key informants

Resonance testing included the following key statements (the researchers varied the ordering of these statements for each key informant interviewed):

1. Low- and middle-income countries (LMICs) are thought to have less and poorer quality data than high-income countries. They are also seen as having less capacity for data use and analytics needed to reach the Sustainable Development Goals (SDGs).
2. Strong data systems can have a significant impact on inequities in health and development.
3. The data produced by LMICs are hardly used domestically and, therefore, there is scant country-level demand to drive a larger supply and better quality of data.
4. Lack of data use stems from a culture where data collectors see data as flowing up but not back down for their use (eg, data going to ministries of health, but not back down to frontline workers who collect the data).
5. Lack of data use stems policies and practices that encourage data silos (localization).
6. Policies and practices encouraging data localization are common in LMICs and limit the extent to which data-informed decision-making is possible.
7. Harmonization of data can lead to significant improvement in health outcomes. However, careful consideration of the environmental context, such as connectivity and user friendliness, is key to the success of building interoperable systems.
8. Data feedback from electronic health systems has improved quality, increased efficiency, and lowered cost of medical care in LMICs.
9. Data collected outside of traditional health management information systems (HMIS), such as cell phone and geospatial data, are needed for anticipating and responding to disease outbreaks, particularly outbreaks in underserved or rural communities.
10. Data and data science are the key to transforming care delivery and achieving health equity in LMICs.