



Travelling numbers and broken loops: A qualitative systematic review on collecting and reporting maternal and neonatal health data in low-and lower-middle income countries

Jil Molenaar^{a,b,*}, Lenka Beňová^a, Aliko Christou^a, Isabelle L. Lange^{c,d}, Josefien van Olmen^b

^a Institute of Tropical Medicine Antwerp, Nationalestraat 155, 2000, Antwerp, Belgium

^b University of Antwerp, Doornstraat 331, 2610, Wilrijk, Belgium

^c London School of Hygiene and Tropical Medicine (LSHTM), London, United Kingdom

^d Center for Global Health, Technical University of Munich (TUM), Munich, Germany

ARTICLE INFO

Keywords:

Maternal health
Neonatal health
Systematic review
Indicators
Routine data
Complex adaptive systems
Health information system

ABSTRACT

Data and indicator estimates are considered vital to document persisting challenges in maternal and newborn health and track progress towards global goals. However, prioritization of standardised, comparable quantitative data can preclude the collection of locally relevant information and pose overwhelming burdens in low-resource settings, with negative effects on the provision of quality of care. A growing body of qualitative studies aims to provide a place-based understanding of the complex processes and human experiences behind the generation and use of maternal and neonatal health data.

We conducted a qualitative systematic review exploring how national or international requirements to collect and report data on maternal and neonatal health indicators are perceived and experienced at the sub-national and country level in low-income and lower-middle income countries. We systematically searched six electronic databases for qualitative and mixed-methods studies published between January 2000 and March 2023. Following screening of 4084 records by four reviewers, 47 publications were included in the review. Data were analysed thematically and synthesised from a Complex Adaptive Systems (CAS) theoretical perspective.

Our findings show maternal and neonatal health data and indicators are not fixed, neutral entities, but rather outcomes of complex processes. Their collection and uptake is influenced by a multitude of system hardware elements (human resources, relevancy and adequacy of tools, infrastructure, and interoperability) and software elements (incentive systems, supervision and feedback, power and social relations, and accountability). When these components are aligned and sufficiently supportive, data and indicators can be used for positive system adaptivity through performance evaluation, prioritization, learning, and advocacy. Yet shortcomings and broken loops between system components can lead to unforeseen emergent behaviors such as blame, fear, and data manipulation. This review highlights the importance of measurement approaches that prioritize local relevance and feasibility, necessitating participatory approaches to define context-specific measurement objectives and strategies.

1. Introduction

Health indicators provide a measurement of a specific health dimension in a given population (PAHO, 2018). Over the past decades, the field of maternal and newborn health has witnessed the development of a large number and wide range of health indicators. The Millennium Development Goals (MDGs) signed in 2000 promoted quantitative measurement to demonstrate persisting burdens and track progress.

Indicators were increasingly embraced as ‘credible, apolitical and authoritative’ tools for understanding what kinds of policies and interventions are most effective (Erikson, 2015, p. 1157). In the years that followed, the focus on indicators intensified through accountability initiatives such as the 2011 Commission on Information and Accountability (CoIA) and the 2015 Global Strategy for Women’s, Children’s, and Adolescents’ Health, as well as the Sustainable Development Goals (SDGs) with their extensive set of targets and indicators (Every Woman

* Corresponding author.

E-mail address: jil.molenaar@uantwerpen.be (J. Molenaar).

<https://doi.org/10.1016/j.ssmph.2024.101668>

Received 12 November 2023; Received in revised form 27 February 2024; Accepted 2 April 2024

Available online 4 April 2024

2352-8273/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

Every Child, 2015; SDG Knowledge Platform, 2015). A vast amount of measurement work has been done specifically on maternal and newborn health, including through technical advisory groups such as Mother and Newborn Information for Tracking Outcomes and Results (MoNITOR). Few maternal and neonatal health indicators are supported across all key monitoring initiatives, and some of the indicators which are commonly promoted across global initiatives – notably Maternal Mortality Ratio (MMR) and Skilled Birth Attendance (SBA) rates – are notoriously hard to measure and benchmark. In addition, there are many ‘aspirational’ indicators for which common definitions and data collection methodologies are not (yet) developed (Moller et al., 2018).

Many maternal and child health monitoring initiatives currently heavily rely on population-based household surveys such as the Demographic and Health Surveys (DHS). Indicators can also be generated by using data directly from routine health information systems that aggregate data from health facilities (Maïga et al., 2019), rather than from non-routine data collection like surveys and health facility-assessments. However, using routine data is not without challenges. Generating statistics from health facility data can be impeded by issues with completeness and accuracy of reporting, as well as with defining target populations used as denominators (Maïga et al., 2019; Nyamtema, 2010).

Balancing local relevance and global applicability in measurement practices can generate tension. Data needs differ across health system levels, ranging from clinical usefulness at the level of care provision to strategic decision-making by national and international policymakers (AbouZahr, Adjei, & Kanchanachitra, 2007), see Fig. 1. It has been argued that the use of standardised quantitative metrics for globally comparable data risks decontextualizing information, hereby facilitating unintended or purposeful ignorance of local circumstances (Geissler, 2013). Particularly in low- and middle-income countries, the focus on generating globally comparable data has also distracted attention and funds from efforts to prioritize the strengthening of national health information systems (Storeng & Béhague, 2017). In order for data to be meaningful and inform local decision-making to improve quality of care, they must be relevant to local burdens of disease, reflect local definitions, and be accessible in comprehensible formats (WHO, 2008). A focus on international standardisation and comparison might limit or overshadow the gathering of information that is locally relevant.

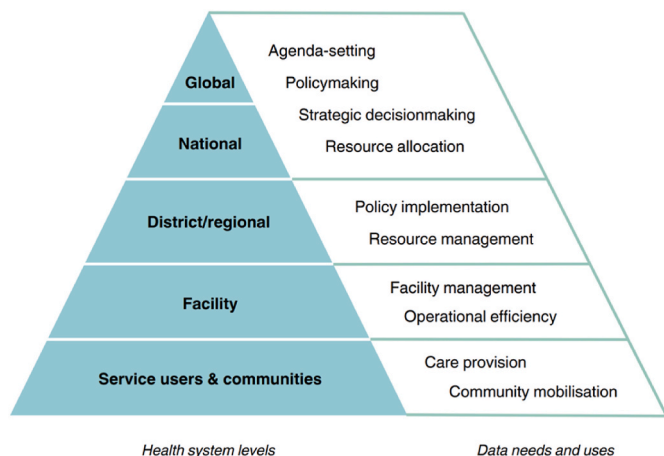


Fig. 1. Pyramid of data needs and uses across health system levels, adapted from AbouZahr et al. (2007).

In addition, top-down demands for measurement and surveillance may pose overwhelming burdens on health workers and other sub-national actors, particularly in contexts of time and resource scarcity (Strong, 2020).

There is a growing interest in interrogating what lies ‘behind the measures’: exploring the imperfect ways in which phenomena become quantified and recorded, the processes through which data acquire meaning, as well as how they influence individual and systemic behaviours (Brunson & Suh, 2020). Indicators in maternal and neonatal health have been described as ‘black boxes’, because the pathways and activities that shape their production are often obscured (Erikson, 2015). Qualitative researchers increasingly attempt to make these pathways visible, demystifying data ‘by tracing their life course and travels amid and with human and nonhuman actors’ (Biruk, 2018, p. 5). Qualitative research can shed light on the key barriers impeding reliable data collection, aggregation, calculation and reporting of maternal and neonatal indicators, as well as on how health system actors and users view the utility and relevance of these measurement practices at the sub-national level.

This systematic review aims to unpack these perceptions and experiences and synthesise key themes around the collection and reporting of maternal and neonatal health data. It focuses on qualitative findings from diverse low- and lower-middle income countries, where the burden of maternal and neonatal mortality and morbidity remains high and where data are often collected and reported in contexts of scarcity and precarity. As a growing body of qualitative studies has aimed to provide a place-based or ‘emic’ understanding of issues arising from demands for data and evidence in LMICs in the field of maternal and newborn health in the past two decades, bringing these findings together is highly informative. While recent reviews have focused on synthesizing experiences with specific types of measurement tools or initiatives such as home-based records (Joseph et al., 2022) or maternal and perinatal death surveillance (Willcox et al., 2023), this is the first systematic review of qualitative studies exploring how national or international requirements to collect data and report on maternal and neonatal health indicators are perceived and experienced at the sub-national and country level in low-income and lower-middle income countries.

We theorize maternal and neonatal health data from a complex adaptive systems (CAS) perspective. CAS theory analyses the structure, interactions, and dynamics of complex adaptive systems, which change and evolve depending on the dynamic components of which they consist (Turner, Hanel, & Klimek, 2018). In contrast to linear cause and effect inquiry, complex systems thinking underscores the significance of nonlinear causality and contextual elements. We thus conceptualize maternal and neonatal health indicators not as fixed, neutral entities, but rather as the outcome of a dynamic, multi-stage, context-dependent process. Their creation and uptake is influenced by myriad system ‘hardware’ factors such as infrastructure, financing, staffing and supplies, as well as system ‘software’ factors including tangible (e.g. leadership, designated roles, rules and procedures) and intangible (e.g. power differentials, values and norms) elements (Asefa, McPake, Langer, Bohren, & Morgan, 2020). In line with CAS thinking, this review aimed to identify key system characteristics shaping data collection and reporting of maternal and neonatal health indicators and identify instances of system interactions, feedback loops and adaptivity.

2. Methods

2.1. Overall approach

We conducted a systematic review of qualitative studies aiming to understand how national or international requirements to collect and

report data to produce maternal and neonatal health indicators are perceived and experienced at the sub-national and country level in low-income and lower-middle income countries. The protocol was registered on PROSPERO in February 2023 (ID: CRD42023400587). Our methods drew from and combined elements of various documented methods for synthesis of qualitative research. In line with the objective of meta-ethnography as proposed by Noblit and Hare (1988), we aimed to draw together insights from qualitative work to facilitate an interpretation that is more complete in depth and breadth than the interpretations from the single studies. Following the key characteristics of qualitative research synthesis outlined by Sandelowski and Barroso (2007, p. 22), we pursued a systematic and comprehensive retrieval of publications reporting on relevant qualitative studies, systematically employed qualitative methods to integrate the findings from these reports, and aimed to reflexively report the procedural and interpretive moves made over the course of the review process. Our synthesis was also informed by the meta-ethnographic approach outlined by Toyé et al. (2014). This approach distinguishes between first-order constructs (participants’ experiences, narratives, and interpretations) and second-order constructs (the primary authors’ interpretations). Through our review, we further abstracted first- and second-order constructs to develop third-order constructs: our interpretations of the participants’ and original authors’ interpretations.

We deliberately took a flexible approach in setting the methodological requirements for the studies included in our review. It has been argued that a qualitative research synthesis should attempt to ensure that included studies do not ‘co-opt’ the label qualitative research, and reviews should exclude ‘studies that, albeit informative, offer no more than another ‘surface understanding’ of human experience’ (Sandelowski & Barroso, 2007, p. 40). However, we opted to include all studies reporting qualitative findings on the topic of interest, including qualitative findings from mixed-methods studies (see Table 1). We intentionally set out to synthesise a body of studies diverse in terms of topics, styles and methods, and accepted from the outset that some included studies would offer considerably more depth, nuance and reflexivity than others. Although papers providing rich accounts and carefully elaborated interpretations feature more heavily in the final analysis and write-up, we still considered it valuable to include all relevant studies in order to represent voices and perspectives from a wider range of contexts.

Table 1
Inclusion and exclusion criteria.

	Inclusion criteria	Exclusion criteria
Study design and methods	Studies reporting qualitative empirical findings. Mixed-methods studies are included if qualitative findings are analysed and reported in the publication.	Studies reporting solely quantitative data, as well as commentary-style articles which do not report on empirical data.
Topic of interest	Studies reporting experiences with or perceptions of data collection and/or reporting relevant to the health outcomes and health services offered to women and neonates during pregnancy, childbirth, the postnatal and neonatal periods. This includes data collection and reporting relating to indicators on service access and availability; service quality and safety; service coverage; risk factors and behaviours; health status; mortality; fertility; health work force; health financing; and supply chain.	Studies reporting on experiences with or perceptions of: <ul style="list-style-type: none"> • Specific measurement devices (e.g. fetal heart rate monitoring devices) • Specific e-health applications (<i>unless</i> the focus is on data collection and reporting)
Population of interest	Health system users (i.e. patients) and health system actors (including in clinical care, operational/managerial, and leadership roles) who have personal experience with the topic of interest.	Individuals who are not health system users or actors.
Context	Studies for which data was collected in low-income countries or lower-middle income countries as per the 2022–2023 World Bank classification.	Studies for which data was collected in upper-middle or high-income countries.
Publication type	Peer-reviewed academic publications with full-text availability.	Study protocols, editorials, (systematic) literature reviews, monographs, conference abstracts, and other.
Publication date	Studies published between January 2000 and the date of the search.	Studies published prior to January 2000.
Language	Papers written in English, French, Spanish, Portuguese or Arabic.	Studies written in all other languages.

2.2. Literature search

We searched six electronic databases: Scopus; MEDLINE (Ovid); Anthrosource; Anthropological Index Online; LILACS; and JSTOR. The search terms were linked to five key domains: qualitative research; maternal and newborn health; measurement; perceptions and experiences; and low- and low-middle income countries. See S1 (Search strategy) for more details. Search terms were developed through extensive prior scoping searches and refined in collaboration with a librarian. Searches were limited to full text articles published between January 1, 2000 and March 27, 2023. Although searches were exclusively carried out using English search terms, all relevant papers written in English, French, Spanish, Portuguese or Arabic appearing in search results were considered for inclusion. Trial searches were run in February 2023 and an initial trial sample of 100 papers was screened by three reviewers (JM, LB, AC) to finalise inclusion and exclusion criteria. See Table 1 for more details on inclusion and exclusion criteria.

Final searches in all included databases were run on March 27, 2023. Records were deduplicated in Zotero and subsequently imported in Rayyan. In the first round of screening, four reviewers (JM, LB, AC, ILL) screened titles and abstracts against the set inclusion criteria. For quality control, the first author re-screened (JM) 15% of the records, which was deemed sufficient as conflicting decisions regarding inclusion versus exclusion were very rare (0.3%). A total of 101 records were retrieved for full-text reading, of which 48 were excluded. The most common reasons for exclusion were articles not focusing on experiences and perceptions of data collection or reporting (28 articles) or not reporting qualitative findings (15 articles). Citation searching in retrieved papers yielded five additional papers meeting inclusion criteria. This resulted in 58 articles being selected for preliminary coding in NVivo (see Fig. 2).

2.3. Relevance appraisal

We originally planned to carry out quality assessment of all papers selected for preliminary coding using the Critical Appraisal Skills Programme (CASP) method (CASP, 2023), which involves evaluating various components of the research as they are written up (e.g. methods, ethical considerations). However, after preliminary attempts to use the CASP checklist we found it had limited relevance for our review, as a paper’s CASP score did not necessarily correspond to its value for our qualitative synthesis. Therefore, we agreed to focus on ‘relevance

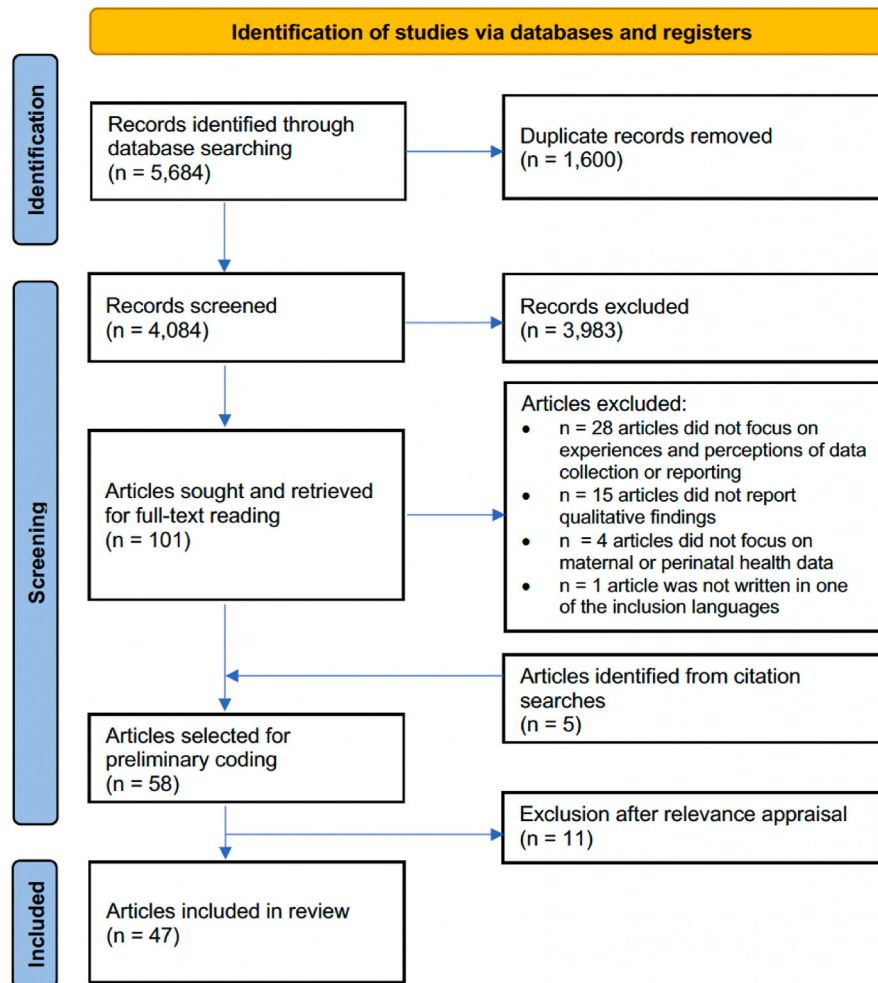


Fig. 2. PRISMA flow diagram selection of studies.

appraisal' instead, which involved excluding papers which did not have sufficient relevant first- and second-order constructs that could be coded in NVivo. As a cut-off, we decided papers that did not have enough relevant information to code at least four fragments or excerpts were to be excluded from the review. This resulted in exclusion of 11 papers which had only tangential relevance to the scope of our review, among which were several mixed-methods papers in which qualitative findings could not be clearly distinguished from quantitative findings. This resulted in the final inclusion of 47 papers.

2.4. Data extraction and analysis

Four authors (JM, JVO, AC, ILL) participated in full-text reading and thematic analysis of the 47 articles included in the review. A data extraction table was completed detailing the main objective of the study, study location, study design, type of participants, health system level focus, summary of main findings, and key analytical themes. Following an exchange of reading notes and discussions among the four authors, a preliminary inductive coding tree was agreed upon consisting of 29 main codes subdivided into 7 descriptive themes. Main codes in the coding tree were key concepts, themes and metaphors in the studies, typically using the phrasing from the article(s) they were identified from. A brief written description or interpretation of each code was agreed upon to clarify what was understood to be covered by that code. For example, the code *parallel systems* was described as 'lack of standardised formats; mismatches between systems; interoperability issues;

double burden of digitised and paper-based systems'. The first author (JM) then used this coding tree to code all selected articles in NVivo (release 1.7.1) in chronological order, starting with the oldest paper. Coding focused on the results and discussion sections of selected articles. Some codes were added or renamed and some descriptions were expanded during the coding process.

Coding in NVivo of the 47 included papers was followed by the creation of thematic tables (see supplement S3) in which each code was illustrated with several example extracts and quotes, providing a type of descriptive 'meta-summary' of the findings (Sandelowski & Barroso, 2007). Subsequently, the first- and second-order constructs (research participants' and original authors' interpretations) summarized in these tables were synthesised to develop third-order constructs. Through several discussions among JM, LB and JVO which revolved around attempting to unite key themes into a common framework, consensus emerged that CAS theory provided a useful lens for higher-order interpretation. In doing this, we focused on findings not yet synthesised elsewhere. For example, barriers and facilitators of health worker motivation have been covered elsewhere (e.g. Shipton, Zahidie, & Rabbani, 2017; Stokes et al., 2016; Willis-Shattuck et al., 2008) and were therefore not prioritised in our analysis. A collaborative approach was thus taken in deciding upon third-order constructs through discussions within the review team, which also served to challenge individual interpretations. The key concepts from CAS theory considered relevant to our findings – notably the distinction between system hardware and software, as well as a focus on system adaptivity and emergent

behaviours – were used to develop a single visualization which we then used to structure the write-up of the findings.

3. Findings

We first summarize descriptive characteristics of the studies included for review (n = 47), including their geographical distribution, publication date, types of participants, and research methods. We then present our findings thematically in three main sections.

3.1. Descriptive characteristics

More than half (28) of the included papers focused on countries in Sub-Saharan Africa, of which 22 were from Eastern and Southern Africa (Table 2). The second most commonly represented World Bank region was South Asia (8 papers), followed by East Asia and Pacific (3 papers) and Latin America and the Caribbean (1 paper). Most papers reported on studies conducted in only one country (39 papers). There was a considerable increase in publications on the topic of interest over time. Although we searched for papers published from 2000 onwards, only four of the included papers were published between 2008 and 2012, 13 papers were published between 2013 and 2017, and 30 papers were published between 2018 and the date of the search.

Over two-thirds (31) of the studies included participants across several geographical and health system levels (i.e. community, facility, district, regional, national, international levels) making it challenging to categorise papers by this characteristic. However, the most frequent group of participants was at the health facility level with 31 studies including participants at this level. Relatively few health system users or patients were represented. As a result, facility-level perspectives and experiences are prominent in the presentation of our findings. For more details on the type of participants, see supplementary material S2. Around one-third (17) of the studies based their results on interviews (semi-structured or in-depth) only, while two studies employed only focus group discussions (FGDs). Fourteen studies combined interviews and FGDs. Ten papers characterized their study as being ethnographic research or as following an ethnographic approach. These were typically studies where observation formed an important part of data collection

and the fieldwork and write-up was guided by a theoretical approach. Four additional studies included observation in their methods, but did not characterize their approach as ethnography. To give a sense of the thematic scope and focus of the studies, we categorized them into five main topics (see S2). Most articles focused on experiences and perceptions of routine data collection and reporting (26 articles). The other main topics were experiences and perceptions in relation to goals/targets and data use (9), digital solutions and e-health applications (5), surveys (4), and home-based records (3).

3.2. Thematic analysis

The publications included for review reported a wealth of experiences and perceptions related to the collection and reporting of data on maternal and neonatal health indicators at the sub-national and country level in low-income and lower-middle income countries. A visual representation of our synthesis is presented in Fig. 3. The figure highlights how data flow and processes related to it take place across health system levels, from the individual to the country and international level. Documentation and reporting are shaped by a variety of contextual drivers and interactions. We distinguish between system hardware and system software issues, which shape means and ability, motivation, and agency at the individual level. In an ideal scenario, data contribute to health system adaptivity and improvement through performance evaluation, prioritization, learning, and advocacy (see blue oval ‘data use’). However, when there are deficiencies in system hardware and software and feedback loops are broken, unintended shifts and emergent behaviors can occur (see red oval).

We present findings in three main sections: 1) system hardware, 2) system software, and 3) system adaptivity & emergent behaviour. We follow the key themes presented in Fig. 3, providing non-exhaustive examples from the included studies. To differentiate between participants’ narratives and experiences (first-order constructs) and primary authors’ interpretations (second-order constructs), we format these differently. First-order constructs, i.e., direct quotes, are presented in *italic* in double quotation marks, while second-order constructs are quoted in single quotation marks.

Table 2
Geographical distribution of included papers.

Region (number of low- and lower-middle income countries in the region)	Number of papers (n = 47)	Countries represented	Paper references
Sub-Saharan Africa (40)			
West & Central (17)	6	Burkina Faso, Cameroon, Ghana, Nigeria, The Gambia	(Gooden et al., 2021; Melberg, Diallo, Storeng, Tylleskär, & Moland, 2018; Ngwakongnwi, Atanga, & Quan, 2014; Rerimoi, Niemann, Lange, & Timæus, 2019; Rothstein et al., 2016; Uzochukwu et al., 2022)
Eastern & Southern (23)	22	Ethiopia, Kenya, Malawi, Mozambique Rwanda, Tanzania, Uganda, Zambia	(Chitama et al., 2011; Haws et al., 2010) (Chiba, Oguttu, & Nakayama, 2012; Danielsen, 2017; Dusabe-Richards et al., 2016; Evans, 2018; Hahn, Wanjala, & Marx, 2013; Kumwenda et al., 2017; Mengesha, Steege, Kea, Theobald, & Datiko, 2018; Wagenaar et al., 2017) (Adane et al., 2021; Cogburn, 2020; Dynes et al., 2023; Estifanos et al., 2022; Freedman et al., 2018; Gladstone et al., 2021; Lange, Nalwadda, Kiguli, & Penn-Kekana, 2021; Melberg, Mirkuzie, Sisay, Sisay, & Moland, 2019; Okello, Gerrets, Zakayo, Molyneux, & Jones, 2018; Regeru, Chikaphupha, Bruce Kumar, Otiso, & Taegtmeier, 2020; Unkels, 2023; Unkels et al., 2022) (Kvernflaten, 2013)
Latin America & the Caribbean (5)	1	Nicaragua	
South Asia (7)	8	Afghanistan, Bangladesh, India, Pakistan	(Begum et al., 2020; Christou et al., 2019; Gautham et al., 2016; Nagarajan, Tripathy, & Goel, 2016; Zakar, Zakar, Mustafa, Jalil, & Fischer, 2018; Scott et al., 2020)
East Asia & Pacific (15)	3	Indonesia, Vietnam	(Das, Newton-Lewis, Khalil, Rajadhyaksha, & Nagpal, 2021; Varley, 2023) (Aiga, Nguyen, Nguyen, Nguyen, & Nguyen, 2016; Målqvist et al., 2008; Faza et al., 2022)
Multiple regions	7	Bangladesh, Benin, Malawi, Nepal, Tanzania, Uganda, undefined	(Hunter, Requejo, Pope, Daelmans, & Murray, 2014; Mahadevan and Broaddus-Shea, 2020; Ruysen et al., 2021; Storeng & Béhague, 2014) (Asefa et al., 2022; Salim et al., 2021; Shamba et al., 2021).

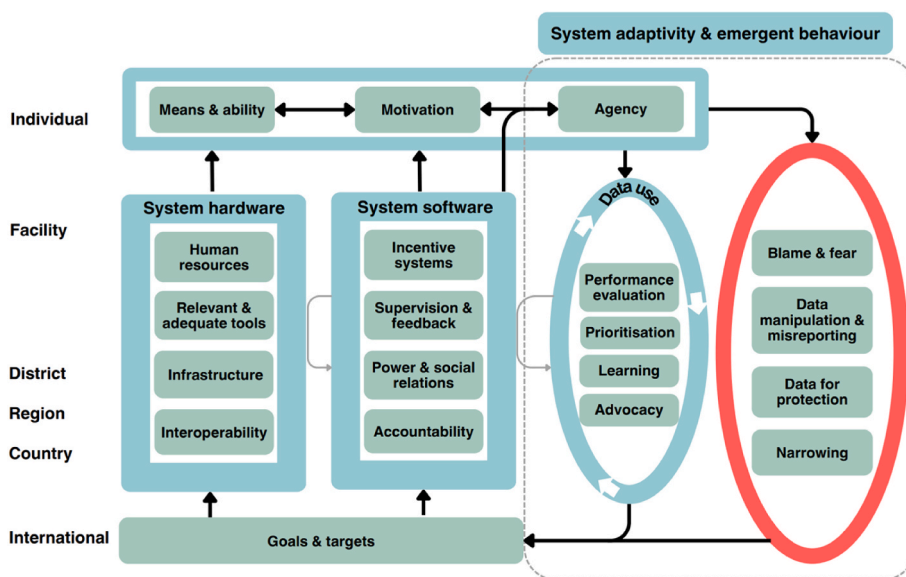


Fig. 3. Maternal and neonatal health data and indicators in complex adaptive systems: a visual representation of key themes.

3.2.1. System hardware

System hardware encompasses fundamental building blocks identified as key elements influencing the process of data reporting and collection. While these components are considered essential prerequisites, they are often inadequate or lacking. Across study contexts, system hardware deficiencies were frequently central to health system actors' accounts of documentation and reporting duties. They were described to combine and interact with system software components to shape experiences at the individual level, as well as to influence data use and other emergent behaviours.

3.2.1.1. Human resources. Human resources are considered a key prerequisite for the collection and reporting of data on maternal and neonatal health indicators. Across included publications, availability of staff is considered a key driver of ability and motivation to document and report. Documentation and reporting duties are typically seen as posing a significant burden on human resources, as they are perceived by health system actors as time-consuming and “very tedious” (Ngwa-kongnwi et al., 2014, p. 33). Considering these tasks must often be performed in clinical settings characterised by heavy workloads and understaffing, demands for data can create tensions at the health facility level:

“Generation of good and quality data is not easy; it is time-consuming. Sometimes the staffs don't have enough time to see patients and they still have to document at the same time”. (Uzochukwu et al., 2022, p. 942)

In challenging work environments, information and data are typically considered secondary to provision of clinical services. A frequently expressed sentiment by maternal and neonatal care workers across settings is that clinical services must take precedence, as expressed by a Prevention of Mother to Child Transmissions (PMTCT) nurse in Malawi: “my duty is to provide PMTCT services, not collecting data” (Kumwenda et al., 2017, p. 308). As a result, reporting is often conducted retrospectively or ‘post-facto’. Many studies reveal a lack of training in documentation and reporting, compounded by high staff turnover necessitating regular retraining in data and documentation skills. At district and regional levels, human resource challenges also limit the uptake of data for monitoring and planning, as it is a labour-intensive task to generate informative reports from large quantities of health data (Nagarajan et al., 2016).

3.2.1.2. Tools. Characteristics of the tools, cards, registers or electronic forms used for documentation and reporting are also a key hardware component. As the most basic prerequisite, tools need to be physically available for the health system actors expected to fill them out. In several studies this was reported to not always be in the case, as noted in a study in Western Kenya:

“[...] the register supply to Bondo was irregular, which forced midwives to use improvised notebooks for three and a half months.” (Chiba et al., 2012, p. 332)

Other characteristics of tools which emerged to be intricately linked to how health system actors perceive and experience data collection and reporting are their ease of use and relevance. A frequent complaint was that health system actors were expected to report data which were not appropriate or relevant for local burdens of disease, available services and diagnostic capacities. When reporting requirements were perceived not to match local realities, this negatively impacted motivation. A key informant in a study on home-based maternal and child health records noted:

“If it is not being used, be ruthless and take it out. Otherwise, you undervalue the importance of the records.” (Mahadevan & Broaddus-Shea, 2020, p. 106)

Mismatches between documentation demands and clinical realities were often described as being discouraging and demotivating. The format and contents of tools and registries used at the health facility level were typically considered to have been designed based on the demands and targets ‘from above’, rather than for healthcare workers' purposes. Definitional and terminology-related issues pose further challenges in documentation and reporting, including inconsistent interpretations of indicators and language discrepancies. For example, Western biomedical categories for pregnancy losses and early neonatal deaths used in surveys and routine data collection formats may not correspond neatly with locally relevant categories. A study in Southern Tanzania found that ‘no Swahili or local language terms existed to differentiate losses by gestational age or vital status at birth, and terms for miscarriage and stillbirth, as well as late-term losses and neonatal deaths, overlapped’ (Haws et al., 2010, p. 1766). If terminology used in documentation and reporting tools does not align with local terms, questions can get lost in translation, potentially resulting in inaccurate data collection. Moreover, the collected data may be less useful to inform local-level change. As such, terminological issues affect both the

perceived relevance of measurement tools and the use of the information that is recorded.

3.2.1.3. Infrastructure. Technical and infrastructure issues also affect system functioning and data flow. Infrastructural challenges may complicate the measurement process itself, leading to issues in data recording. For instance, authors of a study in rural Burkina Faso noted that ‘it was simply impossible to take the blood pressure of the women in antenatal care, the women in labour and the women in family planning clinic at the same time’, as the health facility only had one device to measure blood pressure (Melberg et al., 2018, p. 31). Power outages and internet problems were frequently mentioned, particularly in studies focusing on mHealth or digitised reporting solutions. Absence or non-functionality of devices such as computers, tablets and phones were also identified as key barriers (e.g. Adane et al., 2021; Faza et al., 2022; Nagarajan et al., 2016; Rothstein et al., 2016). Logistical complexities, such as having to send broken tablets for repairs to the capital city (Begum et al., 2020, p. 7), were reported to contribute to delays and inefficiencies. Power cuts were often a major source of worry and a cause of backlog in data entry (Nagarajan et al., 2016, p. 37). Infrastructural challenges “affect timeliness” of data flow “from base levels to higher levels” (Dusabe-Richards et al., 2016, p. 358) and also have a negative impact on motivation.

3.2.1.4. Interoperability. Ideally, data collection tools and systems should be interoperable, meaning data can be unified and aggregated despite originating from different locations and sources. However, lack of interoperability and fragmentation were common issues identified. A study in Uttar Pradesh (India) found that ‘the public sector had not provided any standardized forms for receiving the required data’ on institutional deliveries and ‘facilities used different methods to collect this data’ (Gautham et al., 2016, p. ii41). The absence of standardized reporting formats often resulted in inconsistent documentation and fragmented sharing of information, such as only collecting data ‘on an ad hoc basis when data were required for a service’ (Regeru et al., 2020, p. 340). The existence of ‘parallel systems’ also emerged as a key theme. System actors lamented the presence of multiple overlapping data collection and reporting tools, with indicators duplicated across several reporting forms and registers. A sub-county manager in Kenya was concerned much of this repetition was unnecessary and exacerbated human resource challenges:

“My concern is the issue of duplication of data. I don’t know but I think at the national level, they need to integrate some of these tools. It’s an issue because the health workers are being overwhelmed by the many tools.” (Okello et al., 2018, p. 7)

Similarly, a nurse-midwife in Southern Tanzania pointed out that the burden of having to fill out several registers undermined clinical practice, as “the time you spend on this is more than the care you provide” (Unkels et al., 2022, p. 6).

Parallel reporting to specific NGOs or vertical health programs also contributed to health workers’ reporting burden in some settings (e.g., Regeru et al., 2020). Data demands from international development partners can lead to deprioritization of routine data requested by national or sub-national governments, as reported by Chiba et al. (2012) in Western Kenya, where ‘the provision of maternity registers by an NGO impeded the routine HIS’ (p. 337). In contexts where digitised reporting systems had been rolled out, the burden of parallel systems could also include maintaining both paper and electronic forms in order to be able to cross-check data (Begum et al., 2020) or as a back-up in case electronic data collection systems failed (Unkels et al., 2022).

3.2.2. System software

System software components include tangible and intangible elements relating to human interaction. Key system software factors

emerging from the studies were incentive systems, supervision and feedback, power and social relations, and accountability. These factors interact with system hardware factors to shape experiences at the individual level, as well as the ways in which data are used.

3.2.2.1. Incentive systems. Incentive systems encompass positive incentives and negative sanctions that shape individuals’ perceptions and experiences of data reporting and collecting. Positive incentives for reporting ‘good numbers’ may include both material and non-material incentives and are typically linked to goals and targets coming from the country or international level. In Tanzania’s Dodoma region, Coghburn (2020) observed how nurses working at facilities that performed well in key maternal health indicators ‘would be rewarded with opportunities to attend capacity-building trainings, which according to the District Medical Officer, included the chance to “get some money” in the form of per diems and “do some traveling”’ (p. 5). Non-material incentives also play a significant role, as health system actors strive to be recognized and appreciated for providing good services and for making progress towards goals and targets. Reputational concerns are relevant across different levels of the health system, but may manifest in different ways. At the health facility level, concerns might relate mostly to how the health facility and its staff are perceived by the community and direct supervisors. In Ethiopia’s Oromia and Amhara regions, health care workers and facility managers aspired to ‘rank highly to create a positive image for their facilities compared with other facilities and to gain power or recognition, admiration, appreciation and praise for the facility’ (Estifanos et al., 2022, pp. 4–5). At higher levels, officials might be motivated by comparisons with other countries and the desire to be developing and make progress. Evans (2018) reported that in the Zambian context, ‘international benchmarking and consequent awareness of Zambia’s comparatively poor performance appears to have catalyzed attention relating to maternal health indicators within Ministry of Health’ (p. 231). Numbers are often used for benchmarking and comparison, whether at the subnational level (e.g. between health facilities or regions) or at the international level (between countries). Comparisons can be powerful as they highlight disparities, and “nobody wants to be the worst performer” (Hunter et al., 2014, p. 6). However, benchmarking can also provide strong incentives for data manipulation. Particularly when there are negative sanctions or repercussions in place, such as financial implications or (legal) blame, incentive structures can have a chain of causal effects (see section 3.2.3 System adaptivity & emergent behaviour).

3.2.2.2. Supervision & feedback. Another important element of system software dynamics is the nature of supervision and feedback. Health workers at the health facility level typically considered supportive supervision and feedback as a key driver of motivation to accurately document and report data. However, in most studies, feedback on data reporting from superiors or supervisors was very limited. A nurse in Malawi’s Lilongwe district felt that the lack of feedback from the District Health Officer (DHO) limited the potential for improvement in reporting practices:

“Ah as of me, I have never received feedback from the DHO. However, feedback is useful because sometimes you correct or make things better when you have been told that you did not do well in a particular area, you keep that in mind and, next time, you make sure that the report produced is better.” (Kumwenda et al., 2017, p. 308)

Similarly, a nurse-midwife in a northern state of India lamented that “we never get any feedback on the reports we submit” and as a result, “we never come to know about our mistakes” (Nagarajan et al., 2016, p. 38). In contexts where some level of feedback was provided, supervisory interactions frequently revolved around reviewing the extent of data completeness, rather than the accuracy of the data and their implications for accountability (Das, Newton-Lewis, et al., 2021).

3.2.2.3. Power & social relations. Power hierarchies and social relations are also key components of system software. Supervisory dynamics, as discussed above, are one way in which power dynamics take shape. As demands for data typically come from ‘above’ in hierarchical health systems, they may cause differential pressures and tensions at different levels. Individuals lower in the hierarchy, such as health care users and non-physician healthcare workers, may lack the agency to claim ownership of data and resist unrealistic demands. For example, in a village in Malawi’s central region, health surveillance assistants (HSAs) were expected to produce improved results quickly, and the district health officer explained that if a HSA did not report any progress in their area, they would lose their job. The author noted the power dynamic at play here:

‘With little education, they were easily replaceable. For the HSAs on the other hand, the job was invaluable, as other employment was almost impossible to find.’ (Danielsen, 2017, p. 439)

Power inequities frequently manifested in terms of job insecurity and financial precarity. In many cases, individuals felt they had no choice but to abide by documentation requirements and produce expected results, even if this required data manipulation. Similar to the situation described by Danielsen (2017) in Malawi, an ethnographic study by Das et al. in an eastern Indian state found that managers take advantage of the ‘low social and cultural capital’ of auxiliary nurse midwives (ANMs), who lack formal and informal power to resist demands for data and reporting. One ANM shared:

“What could we do! We cannot leave the job. If the office is always asking for information in the right time and right way, we must quit eating and drinking. And they want to see filled paperwork. We do it in that way. How many works they give; we must find out a way right?” (Das, Newton-Lewis, et al., 2021, p. 7)

It is important to note that health system actors in managerial roles too, in turn, experience pressure and in some cases punitive sanctions from regional, national or international levels. Power dynamics are thus linked with goals and targets and the ways in which accountability to meet them is imposed down hierarchical health systems.

3.2.2.4. Accountability. Closely related to other system software elements, accountability is a fourth key issue. Ideally, data should facilitate a sense of responsibility and agency to learn, consequently driving improvements in health services and outcomes. Data can contribute to creating accountability mechanisms in various directions and at different health system levels: between health workers and health care users and the wider community; among health workers and their superiors; and in relation to targets and goals at higher health system levels. Yet in the studies included for review, the concept of accountability mostly came to the fore in the context of performance targets and evaluation (also see section 3.2.3 System adaptivity & emergent behaviour). Rather than multidirectional accountability, there was a strong focus on accountability upwards, particularly with reference to international goals and donor demands. Based on an ethnographic study exploring the rise of evidence-based advocacy in maternal health, including expert interviews with international-level professionals, Storeng and Béhague (2014) note how some respondents felt uncomfortable with the potential negative impact of imposing accountability demands:

‘Several of those who had participated in the Countdown to 2015 and similar accountability projects demonstrated unsettled ambivalence about the fact that they had contributed not just to agenda-setting but also to the exportation of “target culture” to donor-dependent countries. They noted, for example, that the enforcement of accountability demands can encourage donor recipients to produce fake numbers.’ (p. 271)

Attempts to create accountability mechanisms often result in ‘more or less implicit pressure from higher officials in the health system or political hierarchy’ (Melberg et al., 2019, p. 495). Such tensions are particularly prominent for indicators under close political scrutiny, such as maternal mortality, as pointed out by Kvernflaten (2013) in the Nicaraguan context:

‘The anxiety generated by the pressure to avoid maternal deaths is not easy to deal with, as this is situated within a system of targets and of performance measurement, where government performance also is rewarded or criticised by global actors.’ (p. 38)

This type of ‘anxiety’ can result in unintended consequences and emergent behaviours, as discussed in the next section.

3.2.3. System adaptivity & emergent behaviour

In an ideal case scenario, data are used for health system improvements, positively reinforcing system hardware and software in a cyclical fashion. As summarized in Fig. 3, key types of data use include performance evaluation, prioritization, learning and advocacy. Some studies documented how ‘a state of routine data utilization’ can be used ‘to identify and drive change and share successes’ (Wagenaar et al., 2017, p. 70). Accessibility of timely data across health system levels has also been described as aiding decision-making, and creating ‘stronger links with policy makers for action’ (Mengesha et al., 2018, p. ii78). Supportive leadership was described to be key in promoting a fruitful ‘information culture’ or ‘data-use culture’ (Kumwenda et al., 2017; Wagenaar et al., 2017). Data use for advocacy is documented particularly at higher health system levels, as metrics can be used to lobby national and international stakeholders and hereby mobilise resources (e.g. Hunter et al., 2014). The outcomes of data use can thus be fed back into goals and targets, as well as system hardware and software components.

The included studies also shed light on other types of system adaptivity. When there are systemic deficiencies in hardware and software components, and the loops between system components are broken, a number of emergent behaviours arise. The studies show that a systemic emphasis placed on indicators and performance, when not matched by appropriate and relevant system hardware and software, can create unintended shifts and ripple effects (see red oval in Fig. 3). We discuss four types of emergent behaviours: blame and fear; data manipulation and misreporting; data for protection, and narrowing.

3.2.3.1. Blame and fear. When data flow is not sufficiently supported as the result of systemic deficiencies, study findings across diverse settings suggest performance evaluation often results in environments characterised by blame and fear. Rather than using data for improvement, the focus instead shifts to establishing culpability. Particularly for negative outcomes like perinatal deaths, studies report on the emergence of a ‘culture of blame’ surrounding documentation and reporting (e.g. Christou et al., 2019; Estifanos et al., 2022). Based on interviews with health system actors in Nicaragua, Kvernflaten (2013) noted that ‘the fear of being held responsible for a maternal death was revealed at all levels’, quoting a representative from a women’s organization: “We shiver at the thought of a maternal death in our province” (p. 36). This fear often leads to a pattern of blame-shifting and scape-goating, in which individuals or groups are singled out as responsible. As noted by Cogburn (2020), this draws the focus away from systemic hardware and software challenges:

‘[...] indicator-based care can sometimes cause responsibility, and blame, to fall on the individuals enacting and receiving care, rather than addressing the structural barriers of the system that undermine people’s ability to give and receive care.’ (Cogburn, 2020, p. 7)

Similarly, Adane et al. (2021) found that the culture of blame in the Ethiopian routine health management information results in staff being fixated on not drawing ‘negative attention’ to themselves, ‘which does

not predispose them to raise awareness of systemic weaknesses or help develop genuine accountability' (p. 8). Dynamics of blame and fear were identified across health system levels, as a district-level (woreda) health bureaucrat in Ethiopia explained in relation to maternal deaths:

"Workers or professionals that reported these maternal deaths, they might see that they are going to be blamed by the higher positions ... The health extension workers might expect that they are going to be blamed by the health centre or the health office (woreda level). We, the health office (woreda level) coordinator or officers assume that we are going to get blamed by the higher region, that's why reports on maternal deaths are not here in our hands." (Melberg et al., 2019, p. 495)

Linking to power and social relations, blame and fear thus follow the top-down direction of goals and targets and the ways in which they are enforced along hierarchical health system levels.

A culture of blame and fear can alter incentive systems and create fertile ground for data manipulation and misreporting.

3.2.3.2. Data manipulation and misreporting. A widely discussed emergent behaviour across study contexts was data manipulation and misreporting. This was a sensitive topic, and health system actors would rarely admit to manipulating data themselves, but frequently conceded they had observed it or heard about instances in which it happened. Misreporting encompasses data fabrication, also referred to as "cooking data" (Regeru et al., 2020; Unkels et al., 2023) or "going to the laboratory to make up numbers" (Melberg et al., 2018). Examples of data fabrication include producing numbers on laboratory tests that were never performed, and filling partographs retrospectively including recording of measures which were not actually carried out (Das, Newton-Lewis, et al., 2021; Melberg et al., 2018; Unkels et al., 2022). In Ethiopia, Melberg et al. (2018) reported on a situation where births that occurred before arrival at a health facility were nonetheless reported as facility-based deliveries:

"When a woman arrived during the night with her recently delivered baby, the midwife on call documented it as if she had observed the baby's vital signs the ten first minutes following the delivery. She noted: 'I have to put something, so I put an Apgar score of 8 (out of 10).'" (Melberg et al., 2018, p. 32).

In addition, studies describe deliberate overreporting or underreporting, typically to meet specific goals and targets, avoid blame, and protect reputations. Christou et al. (2019) report on instances of underreporting of stillbirths in Afghanistan's Kabul province:

'Managers referred to practices such as not documenting deaths, modifying or destroying medical records to avoid any documentation, including intentionally misreporting the death to avoid blame. The problem of under-reporting of stillbirths was described by one respondent as 'endemic' in health facilities across the country, admitting that '... the reality that really the numbers, the real number, is manipulated, it's decreased. These are the issues ...'" (Christou et al., 2019, p. 7)

The extent to which data manipulation and misreporting were presented as an outcome of powerlessness, or rather as an expression of agency, differed across publications. Linking to the issues of job insecurity and other power inequities discussed previously, many health system actors framed data manipulation as something they were practically forced to do if they wanted to keep their jobs and function in a work environment characterised by time and resource scarcity. On the other hand, creative ways of dealing with data also often appeared as an avenue to exert agency in challenging work environments. Ways in which system actors take ownership of data and use it for their own purposes are discussed in more detail in the next section.

3.2.3.3. Data for protection. A key emergent use of data, particularly at lower health system levels, can be summarized as 'data for protection'. Rather than data use for learning and health system improvement, it becomes a key way to manage social relationships and safeguard positions. Data for protection can take various forms, but it often has an element of deflection – data are used to redirect the focus and avoid blame and other negative consequences. At the most basic level, documentation can provide evidence that health system actors are doing their job. As stated by a data clerk in Malawi: *"any work minus documentation, it's no work done"* (Regeru et al., 2020, p. 341). One nurse in Malawi's Lilongwe district echoed this sentiment as *"information not recorded is work not done"*, and another pointed out that by documenting the services she provides her supervisors will know she has worked, so *"PMTCT information that I manage acts as my evidence"* (Kumwenda et al., 2017, p. 308). Data for protection might also relate more specifically to documenting the quality and content of care that is provided. Unkels et al. (2022) found that in their study setting in Southern Tanzania, the partograph was used as evidence of provided care and formed the centrepiece of health workers' 'documentation efforts to feel secure':

'Nurses talked about the importance of using the partograph and archiving it to access when mothers came back with a sick newborn. Partographs were kept together in piles in cupboards after they had been counted for the monthly HMIS report, but participants mused about the need to produce something in writing to reduce problems with the community they were so close to. Apart from social risks of bad outcomes, like being blamed by the community, participants were also afraid of legal risks and written documentation seemed to help them cope with this.' (p. 7)

As such, data can provide bidirectional protection: from accusations 'from below' from the community, as well as from supervisors and health system actors 'from above'. Sometimes data are strategically manipulated to achieve this protection. As Varley (2023) notes in the Pakistani context, healthcare workers had become socialized to safeguard themselves and their closest colleagues by skilfully shaping reports to deflect risks. Varley refers to this as records that are 'doctored' to meet expectations, which 'facilitated providers' professional and reputational self-preservation and afforded tactical benefits for hospital administrators' (p. 25). At regional and national levels, government stakeholders may also strategically use or withhold data to protect reputations and achieve political goals. For instance, Kvernflaten (2013) found that regional NGO representatives in Nicaragua suspected that governmental maternal mortality statistics were incomplete, but they had been banned by governmental officials to analyze data on maternal deaths in the area themselves.

Another type of data for protection relates to reporting on issues and health outcomes that are considered particularly sensitive, such as stillbirths, abortions, or unmarried women giving birth. Recording of such events can result in stigma and other negative implications for health care users and their families. For example, parents and health workers in Pakistan participating in a study by Zakar et al. (2018) explained how stillbirths are seen as a type of disability, which is why participants considered it better for these deaths to be concealed as much as possible:

'This is because no one wants to be stigmatized with any disability, primarily because of its negative implications regarding one's image in society. A common belief is that a woman brings her fortune and misfortune with her to her in-laws.' (p. 4)

As confidentiality of patients was not always maintained in this setting, mothers often did not report stillbirths because they feared stigma and social exclusion (Zakar et al., 2018). In fact, the ways in which the confidentiality of data is approached is intricately related

with the safeguarding of social relationships and power structures. [Unkels et al. \(2022\)](#) noted how maternity care providers (MCPs) in Southern Tanzania lived in close proximity to the communities they served, which meant they were invested in maintaining social relationships with that community, as well as protecting their own position:

‘Protecting their own documentation seemed thus important to reduce social disruption through breaches in confidentiality. On the other hand, MCPs had the power to choose whether to hold back or share their clients’ sensitive health data, which may have contributed to their stand in the community.’ ([Unkels et al., 2022](#), p. 8)

Linked to system software issues of power hierarchies and social relations, data for protection thus has a strong relational element vis-a-vis health care users and communities.

3.2.3.4. Narrowing. The concept of narrowing refers to how documentation and reporting may result in a focus that is more limited in scope. In a general sense, the studies included for review demonstrate how quantitative targets and measurement demands can create selectivity in the maternal and neonatal care provision and health outcomes that are prioritised. For example, [Melberg et al. \(2018\)](#) noted how health workers in rural Burkina Faso directed their focus away from services and outcomes that received less vigilant monitoring from the health district, like stillbirths. Accordingly, ‘knowing what not to know’ is key in measurement and reporting, and data flows are ‘defined by layers of inclusion and exclusion of information’ ([Lange et al., 2021](#), p. 468). By strategically selecting information that is recorded and reported, data can present idealized versions of reality, used to ‘create a new narrative of a maternity care where standards were followed, although [only] on paper’ ([Unkels et al., 2022](#), p. 9). Reinforcing this notion, [Varley \(2023\)](#) terms the results of such reporting practices ‘hospital paperworlds’: datasets that are defensively crafted to portray desired outcomes rather than actual achievements, omitting contextual risks and harms. Such narrowing then feeds back into the system in various ways. Indicator-driven prioritization influences clinical practice, and can hereby have negative consequences for quality of care. For example, [Das, Newton-Lewis, et al. \(2021\)](#) observed how non-recording of low birth weights was detrimental to quality of care, as low birth weights would normally require substantial follow-up with parents and other procedures if protocols were adhered to. At the facility level, ‘new narratives’ or ‘hospital paperworlds’ may prevent in-depth understanding of the problems and their roots, thus obstructing attempts to address them. As argued by [Lange et al. \(2021\)](#), omission of information or other types of selective reporting can result in flawed policy recommendations and the ineffective allocation of funds, in turn reshaping system hardware.

4. Discussion

We systematically reviewed qualitative studies to explore perceptions and experiences of collecting and reporting maternal and neonatal health data in low-income and lower-middle-income countries, synthesizing key themes using a CAS framework. We noted a marked increase in papers on this topic in the past 15 years. While this corresponds to a general increase in scientific publications in the past decades ([Bornmann, Haunschild, & Mutz, 2021](#)), it also indicates the relatively recent growth of the body of qualitative studies exploring the implications of demands for data and evidence in LMICs in the field of maternal and newborn health. More than half (60%) of the included studies were conducted in Sub-Saharan Africa. This should be seen in light of roughly half of the world’s low and lower-middle income countries being found in Sub-Saharan Africa ([World Bank, 2022](#)), as well as the high burden of maternal and neonatal mortality in this region ([UNICEF, 2023](#); [WHO, 2023a](#)).

Using a theoretical lens based on CAS, our findings show how targets and demands for data are negotiated through their interactions with system components. Moving away from conceptualizing maternal and neonatal health data as fixed, neutral entities, CAS thinking highlights how they are constructed through interactions between actors and system components spanning geographical levels ([Paina & Peters, 2012](#)). Maternal and neonatal health data are generated by people about people and in turn shape people’s behaviour, resulting in bidirectional dynamics or feedback loops. It is useful to distinguish between system hardware and system software, both of which influence these dynamics and hereby shape people’s perceptions and experiences. Our findings suggest that key system hardware characteristics relate to human resources, relevancy and adequacy of tools, infrastructure, and interoperability, while critical system software elements include incentive systems, supervision and feedback, power and social relations, and accountability. System hardware and software permitting, data and indicators can result in positive system adaptivity by facilitating performance evaluation, prioritization, learning, and advocacy. Yet, when there are shortcomings and disconnects between system components, unforeseen changes and emergent behaviours can arise. Building upon the four types of emergent behaviour described in the findings section, we discuss several types of feedback loops and system interactions in more depth, situating them in relevant bodies of literature.

One key feedback loop concerns the connection between data use and the perceived relevance and feasibility of measurement and data collection. Feasibility might be hampered by weaknesses in system hardware components like human resources and infrastructure, while low perceived relevance might result from mismatches between reporting requirements and local disease burdens, diagnostic capacities, and available services. Studies described how low relevance and feasibility dampen motivation to collect and report data at the individual level, which negatively impacts data quality, which in turn diminishes the relevance of the data to inform services and decision-making.

This is consistent with findings from a growing body of studies exploring sub-national use of routine health data in low-resource settings, which highlight how data-driven quality improvement requires feedback loops involving constructive feedback and a local sense of data ownership ([Lee et al., 2021](#); [Wagenaar et al., 2017](#)). If such feedback loops are not in place, data are often decoupled from their purpose and collected simply because of expectations from above, resulting in empty compliance ([Das, Newton-Lewis, et al., 2021](#); [Mukinda, Van Belle, George, & Schneider, 2020](#)). This resonates with findings from a recent qualitative systematic review on maternal and perinatal death surveillance, which noted how ‘a vicious cycle of under-reporting, inaccurate data, and inadequate review and recommendations’ resulted in demotivation and disengagement ([Willcox et al., 2023](#), p. 62). In her ethnographic account of a maternity ward in Tanzania, Strong notes that a systemic emphasis on data and documentation did not result in ‘deep compliance’, but rather in maternity care providers focusing on ‘surface-level documentation’ to satisfy bureaucratic demands ([Strong, 2020](#), p. 97). As the findings from this review highlight, empty compliance can result in a type of narrowing where data present an idealized version of reality. Such ‘parallel realities’ or ‘paper maternities’ based on documented care may deviate significantly from the care women and their babies actually receive ([Olivier de Sardan, Diarra, & Moha, 2017](#); [Strong, 2020](#)). Numerical data can hereby take on a life of their own and conceal the underlying causes of the problems they were intended to document ([Wendland, 2018](#)).

Another feedback loop concerns the bidirectional relationship between data and power. Maternal and neonatal care is provided in hierarchical health systems characterised by both formal and informal power differentials ([Reddy et al., 2022](#); [Sripad et al., 2022](#)). These hierarchies and social relations shape documentation and reporting practices, and these practices can in turn (re)configure power dynamics. Our findings suggest that health workers frequently feel they have limited options but to conform to documentation demands and produce

expected outcomes, even if achieving this involved manipulating data. This was particularly the case for individuals lower in the power hierarchy who are more likely to face employment insecurity, financial precarity, and gender-based discrimination in many settings (Das, Ramani, et al., 2021; Sripad et al., 2018). This echoes findings from a qualitative synthesis of studies in high-income settings, where midwives explained how strong organizational hierarchies can make them feel like small pawns moved by systemic forces, powerless to change prevailing practices (McFarland, Jones, Luchsinger, Kissler, & Smith, 2020). The way responsibility to collect and report data on maternal and neonatal health indicators is distributed and enforced thus reflects and often reinforces existing power differentials.

On the other hand, the studies also show how data may be used as a means to exercise power and influence. Data might be strategically handled and manipulated to manage social relationships, navigate community norms, and safeguard positions, leading to the creation of 'new power loci' in health facilities and communities (Das, Newton-Lewis, et al., 2021). Rather than authoritative power (power over others), these are typically expressions of discretionary power involving negotiation, sensemaking, and finding agency (Lehmann & Gilson, 2013; Veneklasen, Miller, Budlender, & Clark, 2002). It is a creative type of power that emerges spontaneously in response to systemic conditions and constraints. Following Kingori and Gerrets (2016) interpretation of Scott's (1985) concept of 'weapons of the weak', creative ways of dealing with data may empower health system actors to continue facing their difficult work environment. In some ways, data may act as a weapon or tool for health system actors to express their agency, even if they hold positions with limited influence within organizational hierarchies (Kingori & Gerrets, 2016). The findings of this review thus underline the inherently dynamic and relational nature of the systems in which health indicators are constructed (Duclos, Faye, Ndoye, & Penn-Kekana, 2019; Kielmann, Hutchinson, & MacGregor, 2022). Long before they appear in the form of indicator estimates, data have 'social lives' along their assembly lines: they 'reflect and cohere the social worlds from which they emerge' (Biruk, 2018, p. 200).

Our review also shows how some emergent behaviours, such as data manipulation and misreporting, can become normalised. When deficiencies and disconnects in system hardware and software are chronic, such behaviours may become part of a new systemic equilibrium or 'a habituated component of the organizational culture' (Das, Newton-Lewis, et al., 2021, p. 8). In the case of data manipulation, many health care workers were aware of its ethical implications and indicated they felt discomforted or ashamed. However, it could be hard to behave differently 'in the context of a system that encouraged and rewarded falsification' (Estifanos et al., 2022, p. 5). When behaviours are engrained and act to maintain internal stability in the system, it becomes harder to resist and change them, creating a sense of system inurmountability.

Shortcomings and broken loops in system hardware and software also result in other types of emergent equilibria in organizational cultures. Our review highlights how a data culture allowing data use for positive system adaptivity is dependent on a complex dynamic of supportive system hardware and software elements. When these are not in place, demands for data and documentation often result in a culture of blame instead. Fear of blame in relation to adverse outcomes like perinatal deaths has been documented to be pervasive across diverse contexts, impeding both accurate reporting and data use (Blencowe, Calvert, Lawn, Cousens, & Campbell, 2016; Boyi Hounsou et al., 2022; Cetin, Worku, Demtse, Melberg, & Miljeteig, 2022). Accountability dynamics are altered significantly by fear of blame. Ideally, data contribute to fostering multi-directional accountability, with the ultimate goal of improving health outcomes and experiences at the individual level. Yet, our review shows how system equilibria are typically geared towards upward accountability, in what has been labelled 'target culture' (Storeng & Béhague, 2014). Targets are enforced downwards, and numbers travel upwards. The direction of data flow and accountability

thus mirror and reinforce the frequently critiqued verticality of the global health system (Austveg, 2011; Roalkvam & McNeill, 2016). Strong (2020) has argued how demands for data collection, emanating from a powerful global system, can lead to an accounting culture replacing a caring culture at the health facility level (p. 200). Accordingly, Roalkvam and McNeill (2016) note 'the direction of accountability in the system tends to be reversed, and the individual woman and child, as persons, lost' (p. 84). It is important to note that although such unintended consequences may become deeply engrained, systems are dynamic and equilibria can be shifted. Some optimism is justified – the collection of data on maternal and neonatal health indicators is designed and carried out by people and can therefore be re-imagined and changed.

Our findings underline that improving data flow for positive adaptivity cannot occur in isolation, as it is dependent on people's interactions with system hardware and software elements. This review hereby contributes to the renewed attention towards theorizing the human experience behind maternal and neonatal health indicators. Our findings highlight the importance of qualitative research in nuancing and challenging the information put forward by hegemonic quantitative measures. Such qualitative work reminds us not to fetishize indicator estimates and encourages us to interpret the 'partial stories' they tell (Adams, 2016) with healthy scepticism. More importantly, the experiences and perceptions synthesised in this review also provide lessons on how things can be done differently. By no means do we suggest to stop the collection of routine quantitative health data and the construction of health indicators for monitoring health outcomes and provision of services. On the contrary, we believe investments in routine health information systems in low- and lower-middle income countries are key to ensure the availability of relevant, timely data to inform national and sub-national priorities.

However, our findings do indicate that measurement approaches should give greater priority to place-based relevance and feasibility to allow for productive data collection and use. Greater efforts should be made to avoid that low- and lower-middle income countries have to rely only on information generated through measurement approaches designed far from where they are used, and which are inadequately supported by system hardware and software. Data needs and uses inevitably vary across health system levels, as illustrated in Fig. 1. Health workers use data in clinical contexts; managers require data for operational purposes, and planners and policymakers need data for strategic decisionmaking (AbouZahr et al., 2007). The volume and variety of required data decreases as you move up health system levels, as symbolized by Fig. 1's pyramidal shape. At the very top of the pyramid, only a core set of indicators is needed for monitoring purposes. Yet, at lower health system levels, a larger volume of context-specific data from a variety of sources should guide health service provision and respond to community needs. It may be feasible to include a small number of standard indicators even at these lowest levels. However, the findings of our review underline how local needs should be prioritised in data collection efforts at lower health system levels, rather than taking top-down targets and measurement demands as a point of departure.

This requires a move away from comparability as an overarching aim, as well as an increased openness to qualitative, 'slow' methods such as patient narratives (Adams, Burke, & Whitmarsh, 2014; Oluoch et al., 2023). A data culture centered around local relevance and feasibility requires consensus of what the focus and objectives of measurement should be in a particular context, necessitating a participatory approach (Unkels, 2023). Arenth et al. (2017) propose several dimensions this participatory process should address, including awareness of the need for data use, access to data, skills and knowledge for effective data use, and motivation and opportunity to take data-driven action. International partners and donors also have a responsibility to not 'push for success' (Lange et al., 2021), but rather promote measurement and learning environments that value context-sensitivity and continuous adaptation.

Our review has several limitations. Firstly, we acknowledge that some of the complexity and context-specificity of the human experiences captured in single qualitative studies is unavoidably lost when a diverse body of qualitative studies is synthesised (Sandelowski & Barroso, 2007). We included a relatively large number of studies from highly diverse contexts, and some of the findings presented might not be relevant across all of those contexts. In addition, some contextually specific or unique findings are not represented as priority was given to themes that were shared across studies. We set out to synthesise experiences of both health system users (i.e. patients) and health system actors (including in clinical care, operational/managerial, and leadership roles), yet relatively few health system user perspectives were represented in the included studies (see S2 for participant characteristics). Naturally, this is linked to the fact that most women receiving maternal or neonatal care do not have direct experience with data collection and reporting, with the notable exceptions of home-based records and household surveys. As our search terms did not include specific terms relating to surveys, it is possible we missed some relevant qualitative studies in this domain. The limited representation of health care users' voices in the literature poses a limitation of the review and points to a broader lack of service users' perspectives in our understanding of health indicators. Future research should address this gap, linking with ongoing efforts to promote community participation and engagement in healthcare programmes, e.g. the whole-of-society Primary Health Care (PHC) movement (WHO, 2023b). Our review also identified relatively few studies focusing on national-level perspectives. It seems primary research and/or the publication of such research among national-level actors on this topic is more limited, which may be linked to challenges in accessing these stakeholders for research and the potentially politically sensitive nature of the topic.

Our review included studies with highly disparate methodological approaches, and papers differed significantly in the extent to which they provided in-depth accounts and carefully elaborated interpretations. Many studies relied exclusively on FGDs and interviews, which may not fully capture complex, engrained dynamics surrounding data collection and reporting. We noticed that although system hardware factors were typically extensively covered in papers based on FGDs and interviews only, system software factors and system adaptivity were less well-described in these papers. As Das, Newton-Lewis, et al. (2021) have pointed out, such non-articulated behaviours and social relations might only be adequately understood through ethnographic immersion and observation. Consequently, we call for the pursuit of more high-quality qualitative research on this topic including ethnographic observation, repeated interviews and triangulation of sources whenever feasible. The exclusion of monographs, which typically do combine such methods, poses a limitation. However, efforts were made to link to relevant published monographs in the Introduction and Discussion sections. Finally, we recognize that our positionality as people and researchers – notably as women working at high-income country institutions with pre-existing assumptions and opinions about measurement in the maternal and neonatal health arena – has influenced our analysis.

5. Conclusion

This qualitative systematic review synthesised a heterogeneous body of studies exploring how national and international requirements to collect and report data on maternal and neonatal health indicators are perceived and experienced at the sub-national and country level in low-income and lower-middle income countries. Our findings, synthesised using a CAS theoretical lens, show that maternal and neonatal health data emerge from complex system interactions across health system levels. In discussing factors shaping these dynamics, we distinguish between system hardware (human resources, relevancy and adequacy of tools, infrastructure, and interoperability) and system software elements (incentive systems, supervision and feedback, power and social relations, and accountability). When these system elements align, data and

indicators can promote positive system adaptivity by supporting performance evaluation, prioritization, learning, and advocacy. However, shortcomings and disconnects between these components can lead to unforeseen emergent behaviours. Our findings emphasize the need for measurement approaches that prioritize local relevance and feasibility. This shift should embrace qualitative approaches and involve a participatory approach to define measurement objectives within specific contexts.

Funding

JM was supported by a PhD fellowship from the Flanders research foundation (FWO), reference number 1130323 N. AC was supported by a Postdoctoral fellowship funded by the FWO (reference number 1294322 N). The funding organization had no role in the study design; in the collection, analysis and interpretation of data; in the writing of the article; or in the decision to submit it for publication.

Ethical statement

As this is a systematic review of published literature, no institutional ethics approval was required.

CRediT authorship contribution statement

Jil Molenaar: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Lenka Beňová:** Data curation, Formal analysis, Methodology, Visualization, Writing – review & editing. **Aliki Christou:** Data curation, Formal analysis, Methodology, Writing – review & editing. **Isabelle L. Lange:** Formal analysis, Methodology, Writing – review & editing, Data curation. **Josefien van Olmen:** Formal analysis, Methodology, Writing – review & editing.

Data availability

No data was used for the research described in the article.

Acknowledgements

We would like to thank Barbara Lajeune for her guidance on the systematic search and assistance in developing the search terms.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ssmph.2024.101668>.

References

- AbouZahr, C., Adjei, S., & Kanchanachitra, C. (2007). From data to policy: Good practices and cautionary tales. *The Lancet*, 369(9566), 1039–1046. [https://doi.org/10.1016/S0140-6736\(07\)60463-2](https://doi.org/10.1016/S0140-6736(07)60463-2)
- Adams, V. (2016). Introduction. In *Metrics: What counts in global health* (pp. 1–17). Duke University Press. <https://doi.org/10.1215/9780822374480>.
- Adams, V., Burke, N. J., & Whitmarsh, I. (2014). Slow research: Thoughts for a movement in global health. *Medical Anthropology*, 33(3), 179–197. <https://doi.org/10.1080/01459740.2013.858335>
- Adane, A., Adege, T. M., Ahmed, M. M., Anteneh, H. A., Ayalew, E. S., Berhanu, D., et al., Operational Research and Coaching for Analysts (ORCA)- participants & team. (2021). Exploring data quality and use of the routine health information system in Ethiopia: A mixed-methods study. *BMJ Open*, 11(12), Article e050356. <https://doi.org/10.1136/bmjopen-2021-050356>. Scopus.
- Aiga, H., Nguyen, V. D., Nguyen, C. D., Nguyen, T. T. T., & Nguyen, L. T. P. (2016). Knowledge, attitude and practices: Assessing maternal and child health care handbook intervention in Vietnam. *BMC Public Health*, 16(1), Article Scopus. <https://doi.org/10.1186/s12889-016-2788-4>
- Arenth, B., Bennett, A., Bernadotte, C., Carnahan, E., Dube, M., Thompson, J., et al. (2017). *Defining and building a data use culture*. PATH. https://bidinitiative.org/wp-content/uploads/PATH_Building-Data-Use-Culture_R1.pdf.

- Asefa, A., Dossou, J.-P., Hanson, C., Hounsou, C. B., Namazzi, G., Meja, S., et al. (2022). Methodological reflections on health system-oriented assessment of maternity care in 16 hospitals in sub-Saharan Africa: An embedded case study. *Health Policy and Planning*, 37(10), 1257–1266. <https://doi.org/10.1093/heapol/czac078>. Ovid MEDLINE(R) <2022>.
- Asefa, A., McPake, B., Langer, A., Bohren, M. A., & Morgan, A. (2020). Imagining maternity care as a complex adaptive system: Understanding health system constraints to the promotion of respectful maternity care. *Sexual and Reproductive Health Matters*, 28(1), 456–474.
- Austveg, B. (2011). Perpetuating power: Some reasons why reproductive health has stalled. *Reproductive Health Matters*, 19(38), 26–34. [https://doi.org/10.1016/S0968-8080\(11\)38583-7](https://doi.org/10.1016/S0968-8080(11)38583-7)
- Begum, T., Khan, S. M., Adamou, B., Ferdous, J., Parvez, M. M., Islam, M. S., et al. (2020). Perceptions and experiences with district health information system software to collect and utilize health data in Bangladesh: A qualitative exploratory study. *BMC Health Services Research*, 20(1). <https://doi.org/10.1186/s12913-020-05322-2>. Scopus.
- Biruk, C. (2018). *Cooking data: Culture and politics in an african research world*. Duke University Press. <https://library.oapen.org/handle/20.500.12657/30253>.
- Blencowe, H., Calvert, P. C., Lawn, J. E., Cousens, S., & Campbell, O. M. R. (2016). Measuring maternal, foetal and neonatal mortality: Challenges and solutions. *Best Practice & Research Clinical Obstetrics & Gynaecology*, 36, 14–29. <https://doi.org/10.1016/j.bpobgyn.2016.05.006>
- Bornmann, L., Haunschild, R., & Mutz, R. (2021). Growth rates of modern science: A latent piecewise growth curve approach to model publication numbers from established and new literature databases. *Humanities and Social Sciences Communications*, 8(1). <https://doi.org/10.1057/s41599-021-00903-w>. Article 1.
- Boyi Hounsou, C., Agossou, M. C. U., Bello, K., Delvaux, T., Benova, L., Vigan Guézojè, A., et al. (2022). “So hard not to feel blamed!”: Assessment of implementation of Benin’s Maternal and Perinatal Death Surveillance and Response strategy from 2016–2018. *International Journal of Gynecology & Obstetrics*, 158(S2), 6–14. <https://doi.org/10.1002/ijgo.14041>
- Brunson, J., & Suh, S. (2020). Behind the measures of maternal and reproductive health: Ethnographic accounts of inventory and intervention. *Social Science & Medicine*, 254, Article 112730. <https://doi.org/10.1016/j.socscimed.2019.112730>
- CASP. (2023). *CASP checklists*. CASP - critical appraisal skills programme. <https://casp-uk.net/casp-tools-checklists/>.
- Cetin, K., Worku, D., Demtse, A., Melberg, A., & Miljetej, I. (2022). “Death audit is a fight” – provider perspectives on the ethics of the Maternal and Perinatal Death Surveillance and Response (MPDSR) system in Ethiopia. *BMC Health Services Research*, 22(1). <https://doi.org/10.1186/s12913-022-08568-0>. Scopus.
- Chiba, Y., Oguttu, M. A., & Nakayama, T. (2012). Quantitative and qualitative verification of data quality in the childbirth registers of two rural district hospitals in Western Kenya. *Midwifery*, 28(3), 329–339. <https://doi.org/10.1016/j.midw.2011.05.005>. Scopus.
- Chitamba, D., Baltussen, R., Ketting, E., Kamazima, S., Nswilla, A., & Mujinja, P. G. M. (2011). From papers to practices: District level priority setting processes and criteria for family planning, maternal, newborn and child health interventions in Tanzania. *BMC Women’s Health*, 11. <https://doi.org/10.1186/1472-6874-11-46>. Scopus.
- Christou, A., Alam, A., Hofiani, S. M. S., Rasooly, M. H., Mubasher, A., Rashidi, M. K., et al. (2019). How community and healthcare provider perceptions, practices and experiences influence reporting, disclosure and data collection on stillbirth: Findings of a qualitative study in Afghanistan. *Social Science & Medicine*, 236, Article Scopus. <https://doi.org/10.1016/j.socscimed.2019.112413>
- Cogburn, M. D. (2020). Homebirth fines and health cards in rural Tanzania: On the push for numbers in maternal health. *Social Science & Medicine*, 254, Article Scopus. <https://doi.org/10.1016/j.socscimed.2019.112508>
- Danielsen, L. (2017). Enforcing ‘progress’: A story of an MDG 5 indicator and maternal health in Malawi. *Development and Change*, 48(3), 429–451. <https://doi.org/10.1111/dech.12307>
- Das, P., Newton-Lewis, T., Khalil, K., Rajadhyaksha, M., & Nagpal, P. (2021). How performance targets can ingrain a culture of ‘performing out’: An ethnography of two Indian primary healthcare facilities. *Social Science & Medicine*, 114489. <https://doi.org/10.1016/j.socscimed.2021.114489>
- Das, P., Ramani, S., Newton-Lewis, T., Nagpal, P., Khalil, K., Gharai, D., et al. (2021). “We are nurses – what can we say?”: Power asymmetries and Auxiliary Nurse Midwives in an Indian state. *Sexual and Reproductive Health Matters*, 29(2), 110–126.
- Duclos, D., Faye, S. L., Ndoye, T., & Penn-Kekana, L. (2019). Envisioning, evaluating and Co-enacting performance in global health interventions: Ethnographic insights from Senegal. *Anthropology in Action*, 26(1), 21–30. <https://doi.org/10.3167/aia.2019.260103>
- Dusabe-Richards, J. N., Tesfaye, H. T., Mekonnen, J., Kea, A., Theobald, S., & Datiko, D. G. (2016). Women health extension workers: Capacities, opportunities and challenges to use eHealth to strengthen equitable health systems in Southern Ethiopia. *Canadian Journal of Public Health/Revue Canadienne de Sante Publique*, 107(4–5), e355–e361.
- Dynes, M. M., Daniel, G. A., Mac, V., Picho, B., Asiimwe, A., Nalutaaya, A., et al. (2023). A qualitative evaluation and conceptual framework on the use of the Birth weight and Age-at-death Boxes for Intervention and Evaluation System (BABIES) matrix for perinatal health in Uganda. *BMC Pregnancy and Childbirth*, 23(1), Scopus. <https://doi.org/10.1186/s12884-023-05402-1>
- Erikson, S. L. (2015). Global health indicators and maternal health futures: The case of Intrauterine Growth Restriction. *Global Public Health*, 10(10), 1157–1171. <https://doi.org/10.1080/17441692.2015.1034155>
- Estifanos, A. S., Gezahegn, R., Keraga, D. W., Kifle, A., Procureur, F., & Hill, Z. (2022). The false reporter will get a praise and the one who reported truth will be discouraged’: A qualitative study on intentional data falsification by frontline maternal and newborn healthcare workers in two regions in Ethiopia. *BMJ Global Health*, 7(4), Article Scopus. <https://doi.org/10.1136/bmjgh-2021-008260>
- Evans, A. (2018). Amplifying accountability by benchmarking results at district and national levels. *Development Policy Review*, 36(2), 221–240. <https://doi.org/10.1111/dpr.12213>. Scopus.
- Every Woman Every Child. (2015). *The global strategy for women’s, children’s and adolescents’ health (2016–2030)*. United Nations Secretary General. https://platform.who.int/docs/default-source/mca-documents/rmncah/global-strategy/ewec-globals-trategyreport-200915.pdf?Status=Master&sfvrsn=b42b6d22_4.
- Faza, A., Rinawan, F. R., Mutyara, K., Purnama, W. G., Ferdian, D., Susanti, A. I., et al. (2022). Posyandu application in Indonesia: From health informatics data quality bridging bottom-up and top-down policy implementation. *Informatics*, 9(4). <https://doi.org/10.3390/informatics9040074>. Scopus.
- Freedman, L. P., Kujawski, S. A., Mbuyita, S., Kuwawenaruwa, A., Kruk, M. E., Ramsey, K., et al. (2018). Eye of the beholder? Observation versus self-report in the measurement of disrespect and abuse during facility-based childbirth. *Reproductive Health Matters*, 26(53), 107–122.
- Gautham, M., Spicer, N., Subharwal, M., Gupta, S., Srivastava, A., Bhattacharyya, S., et al. (2016). District decision-making for health in low-income settings: A qualitative study in Uttar Pradesh, India, on engaging the private health sector in sharing health-related data. *Health Policy and Planning*, 31, ii35–ii46.
- Geissler, P. W. (2013). Public secrets in public health: Knowing not to know while making scientific knowledge. *American Ethnologist*, 40(1), 13–34. <https://doi.org/10.1111/amet.12002>
- Gladstone, M. E., Salim, N., Ogillo, K., Shamba, D., Gore-Langton, G. R., Day, L. T., et al., EN-BIRTH Study Group. (2021). Birthweight measurement processes and perceived value: Qualitative research in one EN-BIRTH study hospital in Tanzania. *BMC Pregnancy and Childbirth*, 21. <https://doi.org/10.1186/s12884-020-03356-2>. Scopus.
- Gooden, T., Gustafsson, L., Lu, F., Rickard, F., Sith, A., Cummins, C., et al. (2021). Facilitating better postnatal care with women-held documents in the Gambia: A mixed-methods study. *BMC Pregnancy and Childbirth*, 21(1). <https://doi.org/10.1186/s12884-021-03902-6>. Scopus.
- Hahn, D., Wanjala, P., & Marx, M. (2013). Where is information quality lost at clinical level? A mixed-method study on information systems and data quality in three urban Kenyan ANC clinics. *Global Health Action*, 6(101496665), Article 21424. <https://doi.org/10.3402/gha.v6i0.21424>. Ovid MEDLINE(R) <2013>.
- Haws, R. A., Mashasi, I., Mrisho, M., Schellenberg, J. A., Darmstadt, G. L., & Winch, P. J. (2010). “These are not good things for other people to know”: How rural Tanzanian women’s experiences of pregnancy loss and early neonatal death may impact survey data quality. *Social Science & Medicine*, 71(10), 1764–1772. <https://doi.org/10.1016/j.socscimed.2010.03.051>
- Hunter, B. M., Requejo, J. H., Pope, I., Daelmans, B., & Murray, S. F. (2014). National health policy-makers’ views on the clarity and utility of Countdown to 2015 country profiles and reports: Findings from two exploratory qualitative studies. *Health Research Policy and Systems*, 12(1). <https://doi.org/10.1186/1478-4505-12-40>. Scopus.
- Joseph, L., Lavis, A., Greenfield, S., Boban, D., Jose, P., Jeemon, P., et al. (2022). A systematic review of home-based records in maternal and child health for improving informational continuity, health outcomes, and perceived usefulness in low and middle-income countries. *PLoS One*, 17(8), Article e0267192. <https://doi.org/10.1371/journal.pone.0267192>
- Kielmann, K., Hutchinson, E., & MacGregor, H. (2022). Health systems performance or performing health systems? Anthropological engagement with health systems research. *Social Science & Medicine*, 300, Article 114838. <https://doi.org/10.1016/j.socscimed.2022.114838>
- Kingori, P., & Gerrets, R. (2016). Morals, morale and motivations in data fabrication: Medical research fieldworkers views and practices in two Sub-Saharan African contexts. *Social Science & Medicine*, 166, 150. <https://doi.org/10.1016/j.socscimed.2016.08.019>, 1982.
- Kumwenda, W. W. I., Kunyenje, G., Gama, J., Chinkonde, J., Martinson, F., Hoffman, I., et al. (2017). Information management in Malawi’s prevention of mother-to-child transmission (PMTCT) program: Health workers’ perspectives. *Malawi Medical Journal*, 29(4), 306–310. <https://doi.org/10.4314/mmj.v29i4.5>. Scopus.
- Kvernflaten, B. (2013). Meeting targets or saving lives: maternal health policy and Millennium Development Goal 5 in Nicaragua. *Reproductive Health Matters*, 21, 32–40.
- Lange, I. L., Nalwadda, C. K., Kiguli, J., & Penn-Kekana, L. (2021). The ambiguity imperative: “Success” in a maternal health program in Uganda. *Medical Anthropology*, 40(5), Article 5.
- Lee, J., Lynch, C. A., Hashiguchi, L. O., Snow, R. W., Herz, N. D., Webster, J., et al. (2021). Interventions to improve district-level routine health data in low-income and middle-income countries: A systematic review. *BMJ Global Health*, 6(6), Article e004223. <https://doi.org/10.1136/bmjgh-2020-004223>
- Lehmann, U., & Gilson, L. (2013). Actor interfaces and practices of power in a community health worker programme: A South African study of unintended policy outcomes. *Health Policy and Planning*, 28(4), 358–366. <https://doi.org/10.1093/heapol/czs066>
- Mahadevan, S., & Broaddus-Shea, E. T. (2020). How should home-based maternal and child health records be implemented? A global framework analysis. *Global Health Science and Practice*, 8(1), 100–113. <https://doi.org/10.9745/GHSP-D-19-00340>. Scopus.
- Maiga, A., Jiwani, S. S., Mutua, M. K., Porth, T. A., Taylor, C. M., Asiki, G., et al. (2019). Generating statistics from health facility data: The state of routine health information systems in Eastern and Southern Africa. *BMJ Global Health*, 4(5), Article e001849. <https://doi.org/10.1136/bmjgh-2019-001849>

- Mållqvist, M., Eriksson, L., Nga, N. T., Fagerland, L. I., Hoa, D. P., Wallin, L., et al. (2008). Unreported births and deaths, a severe obstacle for improved neonatal survival in low-income countries; a population based study. *BMC International Health and Human Rights*, 8. <https://doi.org/10.1186/1472-698X-8-4>. Scopus.
- McFarland, A. K., Jones, J., Luchsinger, J., Kissler, K., & Smith, D. C. (2020). The experiences of midwives in integrated maternity care: A qualitative metasynthesis. *Midwifery*, 80, Article 102544. <https://doi.org/10.1016/j.midw.2019.102544>
- Melberg, A., Diallo, A. H., Storeng, K. T., Tylleskär, T., & Moland, K. M. (2018). Policy, paperwork and 'postographs': Global indicators and maternity care documentation in rural Burkina Faso. *Social Science & Medicine*, 215, 28–35. <https://doi.org/10.1016/j.socscimed.2018.09.001>
- Melberg, A., Mirkuzie, A. H., Sisay, T. A., Sisay, M. M., & Moland, K. M. (2019). "Maternal deaths should simply be 0": Politicization of maternal death reporting and review processes in Ethiopia. *Health Policy and Planning*, 34(7), 492–498. <https://doi.org/10.1093/heapol/czz075>. Scopus.
- Mengesha, W., Steege, R., Kea, A. Z., Theobald, S., & Datiko, D. G. (2018). Can mHealth improve timeliness and quality of health data collected and used by health extension workers in rural Southern Ethiopia? *Journal of Public Health*, 40(suppl_2), ii74–ii86. <https://doi.org/10.1093/pubmed/ffy200>. Ovid MEDLINE(R) <2018>.
- Moller, A.-B., Newby, H., Hanson, C., Morgan, A., Arifeen, S. E., Chou, D., et al. (2018). Measures matter: A scoping review of maternal and newborn indicators. *PLoS One*, 13(10), Article e0204763. <https://doi.org/10.1371/journal.pone.0204763>
- Mukinda, F. K., Van Belle, S., George, A., & Schneider, H. (2020). The crowded space of local accountability for maternal, newborn and child health: A case study of the South African health system. *Health Policy and Planning*, 35(3), 279–290. <https://doi.org/10.1093/heapol/czz162>
- Nagarajan, P., Tripathy, J. P., & Goel, S. (2016). Is mother and child tracking system (MCTS) on the right track? An experience from a northern state of India. *Indian Journal of Public Health*, 60(1), 34–39. <https://doi.org/10.4103/0019-557X.177298>. Scopus.
- Ngwakongwi, E., Atanga, M. B. S., & Quan, H. (2014). Challenges to implementing a national health information system in Cameroon: Perspectives of stakeholders. *Journal of Public Health in Africa*, 5(1). <https://doi.org/10.4081/jphia.2014.322>. Scopus.
- Noblit, G. W., & Hare, R. D. (1988). *Meta-ethnography: synthesizing qualitative studies*. In *11. Qualitative research methods*. Newbury Park, London: SAGE Publications.
- Nyamtema, A. S. (2010). Bridging the gaps in the Health Management Information System in the context of a changing health sector. *BMC Medical Informatics and Decision Making*, 10, 36. <https://doi.org/10.1186/1472-6947-10-36>
- Okello, G., Gerrets, R., Zakayo, S., Molyneux, S., & Jones, C. (2018). "Every day they keep adding new tools but they don't take any away": Producing indicators for intermittent preventive treatment for malaria in pregnancy (IPTp) from routine data in Kenya. *PLoS One*, 13(1), Article e0189699. <https://doi.org/10.1371/journal.pone.0189699>
- Olivier de Sardan, J.-P., Diarra, A., & Moha, M. (2017). Travelling models and the challenge of pragmatic contexts and practical norms: The case of maternal health. *Health Research Policy and Systems*, 15(1), 60. <https://doi.org/10.1186/s12961-017-0213-9>
- Oluoch, D., Molyneux, S., Boga, M., Maluni, J., Murila, F., Jones, C., et al. (2023). Not just surveys and indicators: Narratives capture what really matters for health system strengthening. *Lancet Global Health*, 11(9), e1459–e1463. [https://doi.org/10.1016/S2214-109X\(23\)00281-4](https://doi.org/10.1016/S2214-109X(23)00281-4)
- PAHO. (2018). *Health indicators: Conceptual and operational considerations*. Pan American Health Organization. https://www3.paho.org/hq/index.php?option=com_docman&view=download&category_slug=health-analysis-metrics-evidence-9907&alias=45249-health-indicators-conceptual-operational-considerations-249&Itemid=270&lang=pt.
- Paina, L., & Peters, D. H. (2012). Understanding pathways for scaling up health services through the lens of complex adaptive systems. *Health Policy and Planning*, 27(5), 365–373. <https://doi.org/10.1093/heapol/czr054>
- Reddy, B., Thomas, S., Karachiwala, B., Sadhu, R., Iyer, A., Sen, G., et al. (2022). A scoping review of the impact of organisational factors on providers and related interventions in LMICs: Implications for respectful maternity care. *PLoS Global Public Health*, 2(10), Article e0001134. <https://doi.org/10.1371/journal.pgph.0001134>
- Regeru, R. N., Chikaphupha, K., Bruce Kumar, M., Otiso, L., & Taegtmeier, M. (2020). 'Do you trust those data?'—a mixed-methods study assessing the quality of data reported by community health workers in Kenya and Malawi. *Health Policy and Planning*, 35(3), 334–345. <https://doi.org/10.1093/heapol/czz163>. Scopus.
- Rerimo, A. J., Niemann, J., Lange, I., & Timæus, I. M. (2019). Gambian cultural beliefs, attitudes and discourse on reproductive health and mortality: Implications for data collection in surveys from the interviewer's perspective. *PLoS One*, 14(5). <https://doi.org/10.1371/journal.pone.0216924>. Scopus.
- Roalkvam, S., & McNeill, D. (2016). What counts as progress? The contradictions of global health initiatives. *Forum for Development Studies*, 43(1), 69–88. <https://doi.org/10.1080/08039410.2015.1134645>
- Rothstein, J. D., Jennings, L., Moorthy, A., Yang, F., Gee, L., Romano, K., et al. (2016). Qualitative assessment of the feasibility, usability, and acceptability of a mobile client data app for community-based maternal, neonatal, and child care in rural Ghana. *International Journal of Telemedicine and Applications*, 2016. <https://doi.org/10.1155/2016/2515420>. Scopus.
- Ruysen, H., Rahman, A. E., Gordeev, V. S., Hossain, T., Basnet, O., Shirima, K., et al., EN-BIRTH Study Group. (2021). Electronic data collection for multi-country, hospital-based, clinical observation of maternal and newborn care: EN-BIRTH study experiences. *BMC Pregnancy and Childbirth*, 21. <https://doi.org/10.1186/s12884-020-03426-5>. Scopus.
- Salim, N., Shabani, J., Peven, K., Rahman, Q. S.-U., Kc, A., Shamba, D., et al., EN-BIRTH Study Group. (2021). Kangaroo mother care: EN-BIRTH multi-country validation study. *BMC Pregnancy and Childbirth*, 21. <https://doi.org/10.1186/s12884-020-03423-8>. Scopus.
- Sandelowski, M., & Barroso, J. (2007). *Handbook for synthesizing qualitative research*. Springer Publishing Company.
- Scott, J. C. (1985). *Weapons of the weak: Everyday forms of peasant resistance*. Yale University Press.
- Scott, K., Gupta, S., Williams, E., Arthur, M., Somayajulu, U. V., & Noguchi, L. (2020). 'i can guess the month ... But beyond that, i can't tell' an exploratory qualitative study of health care provider perspectives on gestational age estimation in Rajasthan, India. *BMC Pregnancy and Childbirth*, 20(1). <https://doi.org/10.1186/s12884-020-03201-6>. Scopus.
- SDG Knowledge Platform. (2015). Indicators and a monitoring framework for the sustainable development goals. *Sustainable Development Solution Network*. United Nations <https://sustainabledevelopment.un.org/index.php?page=view&type=400&nr=2013&menu=35>.
- Shamba, D., Day, L. T., Zaman, S. B., Sunny, A. K., Tarimo, M. N., Peven, K., et al., EN-BIRTH Study Group. (2021). Barriers and enablers to routine register data collection for newborns and mothers: EN-BIRTH multi-country validation study. *BMC Pregnancy and Childbirth*, 21(1), 233. <https://doi.org/10.1186/s12884-020-03517-3>
- Shipton, L., Zahidie, A., & Rabbani, F. (2017). Motivating and demotivating factors for community health workers engaged in maternal, newborn and child health programs in low and middle-income countries: A systematic review. *Journal of the College of Physicians and Surgeons Pakistan*, 27(3), 157–165.
- Sripad, P., Merritt, M. W., Kerrigan, D., Abuya, T., Ndwiwa, C., & Warren, C. E. (2022). Determining a trusting environment for maternity care: A framework based on perspectives of women, communities, service providers, and managers in peri-urban Kenya. *Frontiers in Global Women's Health*, 3(101776281), Article 818062. <https://doi.org/10.3389/fgwh.2022.818062>. Ovid MEDLINE(R) PubMed-not-MEDLINE <2022>.
- Sripad, P., Ozawa, S., Merritt, M. W., Jennings, L., Kerrigan, D., Ndwiwa, C., et al. (2018). Exploring meaning and types of trust in maternity care in peri-urban Kenya: A qualitative cross-perspective analysis. *Qualitative Health Research*, 28(2), 305–320. <https://doi.org/10.1177/1049732317723585>. Scopus.
- Stokes, T., Shaw, E. J., Camosso-Stepinovic, J., Imamura, M., Kanguru, L., & Hussein, J. (2016). Barriers and enablers to guideline implementation strategies to improve obstetric care practice in low- and middle-income countries: A systematic review of qualitative evidence. *Implementation Science*, 11(1), 144. <https://doi.org/10.1186/s13012-016-0508-1>
- Storeng, K. T., & Béhague, D. P. (2014). "Playing the numbers game": Evidence-based advocacy and the technocratic narrowing of the safe motherhood initiative. *Medical Anthropology Quarterly*, 28(2), 260–279. <https://doi.org/10.1111/maq.12072>
- Storeng, K. T., & Béhague, D. P. (2017). "Guilty until proven innocent": The contested use of maternal mortality indicators in global health. *Critical Public Health*, 27(2), 163–176. <https://doi.org/10.1080/09581596.2016.1259459>
- Strong, A. (2020). *Documenting death: Maternal mortality and the ethics of care in Tanzania*. University Of California Press. <https://doi.org/10.1525/luminos.93>
- Thurner, S., Hanel, R., & Klimek, P. (2018). *Introduction to the theory of complex systems*. Oxford University Press.
- Toye, F., Seers, K., Allcock, N., Briggs, M., Carr, E., & Barker, K. (2014). Meta-ethnography 25 years on: Challenges and insights for synthesising a large number of qualitative studies. *BMC Medical Research Methodology*, 14(1), 80. <https://doi.org/10.1186/1471-2288-14-80>
- UNICEF. (2023). *Neonatal mortality*. UNICEF DATA. <https://data.unicef.org/topic/child-survival/neonatal-mortality/>.
- Unkels, R. (2023). On people, data and systems: Perspectives on routine health data processing and its digitalization in Tanzania. *Inst für global folkhälsa/Dept of Global Public Health*. <http://openarchive.ki.se/xmlui/handle/10616/48766>.
- Unkels, R., Al-beity, F. M., Julius, Z., Mkumbo, E., Pembe, A. B., Hanson, C., et al. (2022). Understanding maternity care providers' use of data: A qualitative study in southern Tanzania. <https://doi.org/10.1101/2022.10.11.22280938>.
- Unkels, R., Manzi, F., Kapologwe, N. A., Baker, U., Ahmad, A., Nabiev, R., et al. (2023). Feasibility, usability and acceptability of a novel digital hybrid-system for reporting of routine maternal health information in southern Tanzania: A mixed-methods study. *PLoS Global Public Health*, 3(1), Article e0000972. <https://doi.org/10.1371/journal.pgph.0000972>
- Uzochukwu, B., Mirzoev, T., Okeke, C., Hicks, J., Etiaba, E., Obi, U., et al. (2022). Did an intervention programme aimed at strengthening the maternal and child health services in Nigeria improve the completeness of routine health data within the health management information system? *International Journal of Health Policy and Management*, 11(7), 937–946. <https://doi.org/10.34172/ijhpm.2020.226>. Scopus.
- Varley, E. (2023). Hospital paperworlds: Medical (Mis)Reporting and maternal health in northern Pakistan. *Medical Anthropology Quarterly*, 37(1), 23–41. <https://doi.org/10.1111/maq.12737>
- Veneklasen, L., Miller, V., Budlender, D., & Clark, C. (2002). *A new weave of power, people, and politics: The action guide for advocacy and citizen participation*. Bourton-on-Dunsmore. <https://doi.org/10.3362/9781780444208>
- Wagenaar, B. H., Hirschhorn, L. R., Henley, C., Gremu, A., Sindano, N., Chilengi, R., et al. (2017). 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 Services Research*, 17. <https://doi.org/10.1186/s12913-017-2661-x>. Scopus.
- Wendland, C. (2018). Who counts? What counts? Place and the limits of perinatal mortality measures. *AMA Journal of Ethics*, 20(1), 278–287. <https://doi.org/10.1001/journalofethics.2018.20.3.pfor2-1803>. Ovid MEDLINE(R) <2018>.

- WHO. (2008). *Framework and standards for country health information systems* (2nd ed.). World Health Organization. Health Metrics Network.
- WHO. (2023a). *Maternal mortality*. WHO fact sheets. <https://www.who.int/news-room/fact-sheets/detail/maternal-mortality>.
- WHO. (2023b). *Primary health care—WHO fact sheet*. World Health Organization. <https://www.who.int/news-room/fact-sheets/detail/primary-health-care>.
- Willcox, M. L., Okello, I. A., Maidwell-Smith, A., Tura, A. K., van den Akker, T., & Knight, M. (2023). Maternal and perinatal death surveillance and response: A systematic review of qualitative studies. *Bulletin of the World Health Organization*, 101(1), 62–75G. <https://doi.org/10.2471/BLT.22.288703>
- Willis-Shattuck, M., Bidwell, P., Thomas, S., Wyness, L., Blaauw, D., & Ditlopo, P. (2008). Motivation and retention of health workers in developing countries: A systematic review. *BMC Health Services Research*, 8(1), 247. <https://doi.org/10.1186/1472-6963-8-247>
- World Bank. (2022). New World Bank country classifications by income level: 2022–2023. <https://blogs.worldbank.org/opendata/new-world-bank-country-classifications-income-level-2022-2023>.
- Zakar, M. Z., Zakar, R., Mustafa, M., Jalil, A., & Fischer, F. (2018). Underreporting of stillbirths in Pakistan: Perspectives of the parents, community and healthcare providers. *BMC Pregnancy and Childbirth*, 18(1), 302. <https://doi.org/10.1186/s12884-018-1924-9>