The Use of Data and Dashboards to Improve Quality of Neonatal Hospital Care in Kenya

Christiane Hagel

Nuffield Department of Medicine

Green Templeton College

University of Oxford

A thesis submitted in fulfilment of the requirement for the degree of

Doctor of Philosophy in Clinical Medicine

Trinity 2024

One does not simply design a dashboard…

Abstract

The Sustainable Development Goals (SDGs) aim to reduce neonatal mortality to at least 12 deaths per 1,000 live births by 2030, yet significant newborn care challenges persist in low- and middle-income countries (LMICs), casting doubt on whether this target will be met. The World Health Organization's Every Newborn Action Plan highlights the urgent need to reduce newborn mortality and morbidity and use data to improve quality of care. However, research has focused on data availability, systems and technology for tracking these goals. While data can provide evidence-based insights, improving neonatal care requires translating data into actionable insights that enable users to act based on the feedback they receive. Dashboards have become popular tools in healthcare to translate data into condensed information. In high-income settings, dashboards are used as Audit and Feedback (A&F) interventions to provide performance feedback to clinicians and managers. Although extensive literature on performance feedback and dashboards in healthcare exists, little is known about the effectiveness of dashboards for quality of newborn care improvements in LMICs. This study explores how dashboards may enable the use of data for performance feedback at public hospitals and improve the quality of neonatal hospital care in Kenya. Reflecting on two neonatal care initiatives that have implemented feedback on neonatal indicators, the study examines how data, visually presented with dashboards, can be used by hospital staff in newborn care.

A mixed methods approach explored the design and implementation of dashboards, assessing their alignment with user needs and context to make newborn data actionable for quality improvements. The study drew on the Clinical Performance Feedback Intervention Theory (CP-FIT) and user-centred design literature. A scoping review included 18 articles to learn from newborn hospital dashboard developments globally. A quantitative survey with potential users (n=103) measured dashboard usability, in particular, data visualisation interpretability. Implementers (n=22) from Kenya, Malawi, Nigeria, Tanzania, the UK, and the USA, were

interviewed about dashboard development experiences. Users (n=20), including paediatricians, nurses, data and hospital managers at Kenyan hospitals, were interviewed to understand their engagement with dashboards, the impact on daily clinical practices, and contextual factors.

The background literature and scoping review confirmed that the usability of newborn dashboards is not regularly tested, and user experience studies are missing for LMIC settings. Empirical study findings indicate that dashboards fall short in supporting the complex efforts to improve newborn care due to design and contextual challenges. Around 50% of users struggled to interpret key dashboard visuals. Usability scores and user perceptions revealed challenges in using dashboards even when described as 'very easy to use' and 'useful'. End-user interviews found diverse challenges, including the need for more intuitive dashboards, effective training to use them, and improved data quality and systems to ensure trust. Staff shortages exacerbate these challenges, increasing workloads and reducing available time to use feedback for quality improvements. Implementers, on the other hand, reported challenges to integrate dashboards into existing systems and provided additional insights into wider implementation challenges.

This study underscores the importance and need for an iterative, user-centred design strategy that incorporates user and implementer knowledge and contextual influences. Findings demonstrate that thorough user testing and active user engagement are crucial when developing hospital dashboards for quality improvement. Understanding diverse user perspectives and the context should be the foundation for building, tailoring and implementing dashboards. Dashboards adapted to user needs and institutional contexts that provide reassurance in data quality, data analysis, and the underlying information systems can help to build trust in the feedback presented in dashboards and thus promote data-driven quality improvement.

Acknowledgements

This research would not have been possible without the invaluable support of others. To everyone who walked this journey with me, I owe you my deepest gratitude.

The collective support of my supervisory team was essential to the success of this work. Words cannot fully express my gratitude to my primary supervisor Professor Mike English, for his invaluable support, guidance, and empathy throughout this DPhil. His academic advice has been instrumental in shaping and finishing this research, and I am truly thankful for his unwavering commitment. I also extend my heartfelt thanks to both my secondary supervisors, Professor Sassy Molyneux and Dr Chris Paton, for generously sharing their invaluable knowledge and expertise, even while managing significant time differences. I cannot thank them enough.

I sincerely thank all examiners for their insightful and constructive feedback, which helped me reflect on and refine my research. I was fortunate to receive guidance from Professor Philippe Guérin, Professor Trudie Lang, Professor Chrysanthi Papoutsi, and Professor John Powell, whose input was indispensable. I am especially grateful to my final examiners Professor Philippe Guérin and Professor Mainga Hamaluba, whose thoughtful review and feedback at the final stage enriched the final reflection on my work and informed its future direction.

This research would not have been possible without the generous financial and administrative support from my funders and my college. I would like to express my sincere gratitude to the Economic and Social Research Council (ESRC), Grand Union Doctoral Partnership programme, which provided tuition fee support and enabled my data collection. I am especially grateful to my department, the Nuffield Department of Medicine (NDM) and to Green Templeton College (GTC) for additional and essential funding to cover living expenses, conference attendances and educational programmes.

A special thanks goes to my research participants and research collaborators from the KEMRI's Clinical Information Network and NEST360 for their generosity allowing me to assess their dashboards, for patiently sharing their views and answering my questions. This research would not exist without you.

I would like to extend my heartfelt thanks to my team, the Health Systems Collaborative (HSC) and my colleagues from the KEMRI Wellcome Trust for the intellectual discussions and camaraderie throughout this DPhil journey. A special shout out to Jalemba Aluvaala, David Gathara, Timothy Tuti, Joyline Jepkosgei, Conrad Wanyama, Livingstone Mumelo, Alun Davis, Shobhana Nagraj, and Jake McKnight.

Green Templeton College was my base throughout this DPhil, a place where I spent most of my time, found a community, and received meaningful support. A special thanks to Coral Milburn for the statistical feedback and to Allison Franklin for all the welfare and logistical support. I am deeply thankful to the library team, Kirsty Taylor and Richard Turner, for their patience with my extended book loans and for supporting the initiation of the FemTech book collection. Observatory Street became my home, and I am especially thankful to the lodge, maintenance and accommodation team for the support throughout.

I am forever grateful to my family, who have shaped not only my education but also the foundation beneath all of this, my values and how I approach life. A very special thanks goes to my parents and brother, who played a significant role in my move to yet another country to work and study (yes, again!) in Oxford. I stayed "hungry and foolish", I hope!

To my dearest friends, who have become my family: Eva, Cathrin und Moritz – over 20 years of friendship that has never been in doubt. Moritz (h.M. Heile), you are the friend who tells me what I don't want but need to hear, and who shows up when it counts – that means more than words can express. Eva and Cathi, thank you for your love and for keeping our K&R family strong. Your support and honesty keep me grounded in the hard times, and you celebrate the good ones with me wholeheartedly, as real friends do. You are simply the best.

To ‘my tribe’: you are many, and while I cannot name everyone, I would like to honour a few who made a particular difference during this DPhil. Lucrezia Vittoria Viti, your love and support have truly been a lifesaver, and I am forever grateful. My COVID-19 pandemic support group and dear friends Sylvie, Rita, Sandra and Amen, thank you for turning tough times into bearable, even joyful, moments and walks. To my Oxford people, Amy, Beth, Cathrin, Davide, Guido, Kerlijn, Mark, Marlene, Melvin, Ryan, Sandrena, and Yan, thank you for all the laughter and companionship that made this journey so much richer. To my friends across the globe, Sven, Katha, Chadi, Johanna, Seb Simon, and Zee, thanks for always being there even from afar. And finally, to the incredible Oxford FemTech Society and the global FemTech community: serving as President has been an honour, and I am deeply grateful for the shared passion and purpose that unites us. You inspire me every day!

“Deeds, not words”, is what guides me. I believe that actions speak louder, a principle that has shaped both my life and parts of this thesis. Life has its way of testing us, but resilience, perseverance, and keeping the right people close helped me to keep in mind what matters most. This DPhil truly mattered to me, and I am proud to have seen it through – with appreciation to those who stood with me along the way.

I dedicate this thesis to my grandmother, whose dreams of higher education were denied in times when opportunities for women were limited. She encouraged and inspired me to pursue knowledge with passion, purpose and dedication. In honour of her resilience and strength.

Declarations and contributions

Declaration of authorship

I hereby declare that this thesis is my own work. Any contributions from others and sources used are acknowledged where applicable.

Chapter IV contains content from the research protocol “Exploring the development and use of dashboards in neonatal care: A scoping review protocol” [1], that I first authored for publication; OSF Registries, 09/2024. DOI: https://doi.org/10.17605/OSF.IO/EVQZP.

Supervisors contributed with intellectual guidance on the overall research strategy, analysis and presentation:

- Professor Mike English, Health Systems Collaborative, Oxford Centre for Global Health Research, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom
- Professor Sassy Molyneux, Health Systems Collaborative, Oxford Centre for Global Health Research, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom
- Dr Chris Paton, Health Systems Collaborative, Oxford Centre for Global Health Research, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom & School of Computing, University of Otago, New Zealand

Additional advisory guidance and assistance was provided for the following chapters.

- Chapter IV: Eli Harriss (Keyword Search Strategy Feedback), Yingxi Zhao (Second Reviewer)
- Chapter V: Dr Timothy Tuti and Conrad Wanyama (Survey Design Feedback) & Dr Coral Milburn (Statistical Analysis Advisory)
- Chapter VI: Joyline Jepkosgei (Research Assistant for Qualitative Data Collection)

Funding

Economic and Social Research Council (ESRC)

- Grand Union Doctoral Training Partnership (DTP)
  Tuition Fee Studentship and Fieldwork Funding

University of Oxford

- Nuffield Department of Medicine
  Oxford-Wellcome Institutional Strategic Support Fund (ISSF)
- Medical Sciences Division Goodger and Schorstein Research Scholarship
- COVID-19 Scholarship Extension Fund (CSEF)

Green Templeton College

- Competitive Conference Fund
- Learning Grant

Word count: 47383

(including abstract, excluding acknowledgments, declarations, tables of contents, tables, figures, appendices and bibliography)

Signature

Ch. Hagel

Christiane Hagel

Table of Contents

List of Tables

List of Figures

Abbreviations & Glossary

Abbreviations

A&F	Audit and Feedback
AMR	Antimicrobial Resistance
BIS	BORN Information System
BORN	Better Outcomes Registry & Network Ontario
CI	Confidence Interval
CIN	Clinical Information Network
DHIS2	District Health Information Software, version 2
e-A&F	Electronic Audit & Feedback
EHR	Electronic Health Records
HIC	High-Income Country/ Countries
HIS	Health Information System
HMIS	Health Management and Information System
HTML	Hyper Text Markup Language
IT	Information Technology
KPI	Key Performance Indicator
LMICs	Low- and Middle-Income Countries
NBU	Newborn Unit
NHS	National Health Service
NICU	Neonatal Intensive Care Unit
NEST	Newborn Essential Solutions and Technologies
PMTCT	Prevention of Mother to Child Transmissions
QI	Quality Improvement
QoC	Quality of Care
RAG	Red-Amber-Green

RedCap	Research Electronic Data Capture
RCOG	Royal College of Obstetricians and Gynaecologists
UNICEF	United Nations International Children’s Emergency Fund
UK	United Kingdom
USA	United States of America
WHO	World Health Organization
XML	Extensible Markup Language

Glossary

Audit and Feedback (A&F): A process to improve care and motivate professionals to modify clinical practice by systematically reviewing care against set criteria, implementing changes, and monitoring for improvement, providing summaries of clinical performance over a specific period [2].

Dashboard: A tool that displays data visuals by consolidating and presenting key information on a screen, enabling users to monitor, analyse, and make decisions based on data [3].

Key Performance Indicator (KPI): A quantifiable metric tied to organisational goals, used to evaluate performance over time and assess progress against benchmarks. Performance can be defined as the accomplishment of a given task measured against pre-set standards of for example, accuracy, completeness, quality, cost, and speed [4], [5].

Quality of Care (QoC): According to the WHO, the “*degree to which health services for individuals and populations increase the likelihood of desired health outcomes*”, based on evidence-based knowledge. To achieve it, it requires effective, safe, people-centred, timely, equitable, integrated and efficient health services [6].

Scorecard: A tool that displays data visuals, providing strategic (performance) information to evaluate and monitor an organisation or system. It analyses current strategies and compares them with the organisation’s / company's overall objectives [3].

Statement on the Impact of the COVID-19 Pandemic on this Research

I experienced significant disruptions in the progress of this research due to the COVID-19 pandemic and related measures. In accordance with the University of Oxford Medical Sciences Division guideline, I report these impacts here.

i) Background and Research Prior to the Closure of University Buildings

This research started following my matriculation as a DPhil probationary research student (PRS) at the University of Oxford in October 2019. The first cases of COVID-19 were reported to the World Health Organisation on 31 December 2019. In line with the government guidance, the University closed public spaces, offices and issued working from home for staff and students in March 2020. Until that date, I had office desk space at my team's office. The change and restrictions occurred approximately six months after matriculating for my DPhil in Clinical Medicine. Thereafter, the 'work-from-home' guidance remained largely in place (with intermittent lifting and re-implementation) until February 2022, at which point hybrid at-home/ on-site working routines were recommended by the department.

At the time of closure of university buildings, I had completed a research proposal and conducted a literature knowledge synthesis of my research. Research ethics were approved since my project fitted into a larger research proposal of my supervisor. I was in the process of travelling to Kenya to identify immediate priorities for my study at that time, to develop a specific research plan, data collection and interview tools in collaboration with colleagues from the KEMRI Wellcome Trust. All documents and research work permits for Kenya had been approved and flights had been booked for this first part of fieldwork. All research activities as planned at this moment were cancelled due to the pandemic and the restrictions mentioned above.

ii) Plans for Intended Studies and Changes

Prior to my DPhil research, Kenya has been identified as a region with a particularly high burden of neonatal deaths. I had been part of a research ethics application to conduct in-depth qualitative research with fieldwork in Kenya planned to start in 2020. Following my initial literature research during the first months as a PRS, I planned to conduct a qualitative study that involved hospital staff and government officials, including field observations over approximately one year in total. The thesis committee approved this research plan during the first meeting in January 2020.

However, the COVID-19 pandemic significantly impacted this research plan, particularly the data collection and fieldwork. Due to travel restrictions and limitations of in-person fieldwork in Kenya, it became necessary to adapt the original methodology and selected methods from in-depth, face-to-face interviews, focus group discussions, and observations to remotely conducted virtual research. Planned on-site activities, especially those including critical healthcare workers, were either delayed or shifted online. I was forced to reduce the qualitative component and introduced a scoping review and an online survey. Consequently, I had to revise my research proposal and prepare a new ethics application to accommodate the updated methodology. Ethics for the revised proposal was approved through the University's Oxford Tropical Research Ethics Committee (OxTREC) in June 2021. However, field visits in Kenya were only possible in April 2022.

Due to these changes, I had a considerable additional workload and delays which led to a substantial reduction of time available to conduct my DPhil research as originally planned. Finally, the title of my research project was changed from ‘The use of data to improve quality of neonatal hospital care in Kenya and help strengthen health systems” to “The use of data and dashboards to improve quality of neonatal hospital care in Kenya”.

iii) Details from Personal Student COVID-19 log

The COVID-19 pandemic also resulted in changes in my personal and wider professional circumstances: I had to shift from on-site working to working from home for approximately two years, which drastically reduced opportunities for face-to-face contact and in-person discussions with supervisors, colleagues and DPhil peers, as well as researchers in Oxford external to my research group. The restrictions also confined my work environment to my home (for the first six months without sufficient desk space) for a significant period (approximately a year and a half), and personal access to healthcare through the NHS was limited.

iv) COVID-19 Related Research

Not applicable. I did not shift my research focus to a COVID-19-related project because the previously identified research gaps and needs remained critical. The pandemic may have further intensified the need for and importance of addressing these original research needs.

Please note: In this thesis, references to specific sub-sections are denoted using Arabic numerals (for example in the format "see 3.1.1."), while references to previous main chapters are denoted using Roman numerals (for example in the format "see Chapter II.").

Chapter I: Introduction

Neonatal mortality remains a significant challenge in low and middle-income countries (LMICs), where too many babies continue to die from preventable causes [7]. The Sustainable Development Goal (SDG) target 3.2.2 aims to reduce neonatal mortality to no more than 12 deaths per 1,000 live births by 2030. With this milestone now only half a decade away, there is growing uncertainty about whether it will be achieved [8], [9].

To address this challenge, strengthening health systems and improving health services are strategies that the World Health Organization (WHO), in collaboration with Ministries of Health and partners, work on to accelerate progress [7], [10]. As part of the Every Newborn Action Plan (ENAP), the WHO stressed the urgent need for tracking progress and developed the Ending Preventable Maternal Mortality (EPMM) Dashboard, which monitors newborn and maternal health in all countries [11], [12]. Health information systems are acknowledged as one of the core blocks of health systems [13]. They are expected to collect quality data and convert data into reliable information, for decision-making to improve health services and routine clinical work [14], [15], for example through feedback to staff about clinical outcomes and processes [2], [16]. Despite the rapid development of digital data collection solutions in LMICs, little is known about the actual use of data – and even less about dashboards – for performance feedback. Using high-quality data for tracking progress and informing decision-making is not only a critical global and national effort, but also a service management strategy at the facility level. However, local data use is reported to remain limited in LMICs due to challenges such as data quality and inefficiencies of the current data and information systems in place [17], [18].

In Kenya, health information systems are complex, relying on a mix of paper-based and electronic health record (EHR) systems, medical registries, surveys and national health management information systems, which creates a high data collection burden on hospital staff [15], [19]. Efforts are made to use patient data to improve quality of newborn hospital care. The introduction of Audit and Feedback (A&F) interventions using dashboards to provide hospital staff with information on how they perform has shown promising results [20], [21], [22]. For example, public hospitals that participated in clinical networks which integrate clinical performance feedback have improved documentation quality and prescribing accuracy in Kenyan newborn units [23], [24]. Despite these advancements, there remain gaps in understanding exactly how data can be used by healthcare workers on the ground to improve Quality of Care (hereafter: QoC). Understanding how visualised information is interpreted is not only important for enhancing the impact of dashboard quality improvement interventions but also for ensuring the safety of clinical decisions by reducing risks of misinterpretation and incorrect action. While dashboards can make complex data more accessible and easier to comprehend by visualising them in the form of graphs, their effectiveness depends on users accurately interpreting visuals, which is heavily influenced by user-centred design [25], [26].

This thesis aims to explore data visualisations and tools currently used to inform various health systems actors in LMICs, with a specific focus on the Kenyan health system and newborn care. Through the collection of both user and implementer perspectives, the study provides insights into how to design dashboards as quality improvement interventions suitable for complex hospital settings like newborn intensive care units at public hospitals in Kenya. The study shifts focus from data to understanding the user experience of making use of data, investigating how dashboards support data-driven quality improvement efforts and how hospital staff can interpret and act upon the visual information provided. Specifically, this research examines the use of data and dashboards to improve the quality of newborn hospital care, taking advantage of two existing initiatives that developed dashboards and provide feedback on QoC (based on neonatal outcome and process indicators) to hospital staff in Kenya. In this setting, it also explores perceptions on the value of data and dashboard interventions among different hospital professional groups to deliver better care for newborns.

The study design and results may be applied to other feedback interventions in similar settings and more widely across health information systems, especially as these become more advanced in LMICs. The research outputs contribute to building learning health systems, supporting continuous data-driven improvement in health outcomes through effective translation of data into actionable practices.

After this introduction, Chapter II provides an overview of the broader literature that informs this research study, including A&F theories and design principles, followed by the study design and methodology in Chapter III. Chapter IV presents findings from a scoping review synthesising published research on neonatal dashboards, their design, and their applications in newborn care globally. Chapters V and VI present the findings from the empirical, mixed methods study. A survey assessed how different user groups interpret dashboard content (Chapter V), and an in-depth qualitative inquiry adds contextual insights on dashboard use in practice and explored design and implementation strategies (Chapter VI). These findings are synthesised and discussed in Chapter VII to reflect on how data and dashboards can be optimised for performance feedback. Finally, I provide conclusions and brief recommendations in Chapter VIII. How the chapters informed each other can be seen in Figure 1.

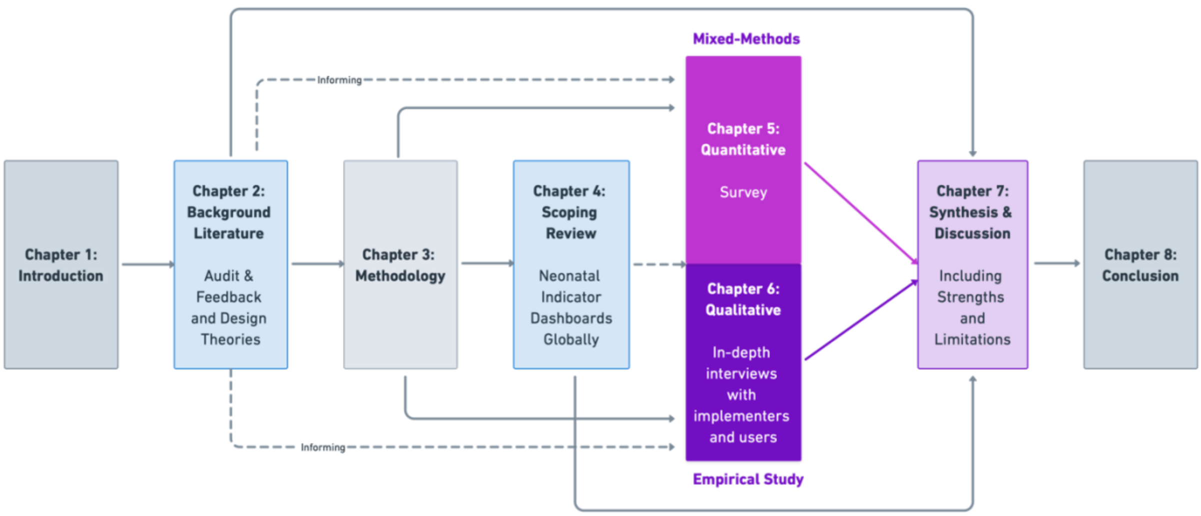


Figure 1: Overview of how thesis chapters inform others and form the discussion and conclusions.

Chapter II: Background

This chapter presents how health systems rely on data to monitor and improve newborn care and to achieve the Sustainable Development Goal (SDG) 3.2., which aims to reduce newborn mortality. I start by exploring quality of newborn hospital care and the role of performance indicators and information systems. Clinical performance feedback and dashboards are highlighted as interventions to use data for QoC improvements. I then present fundamental design principles that are relevant for developing and implementing dashboards in the healthcare setting. The chapter concludes by positioning this research to close literature and practical gaps and presenting the thesis' research questions.

2.1. Health Systems Using Data to Monitor SDG 3.2.

2.1.1. Quality of Newborn Hospital Care in LMICs

In 2022, approximately 2.3 million newborns died globally within the first 28 days of life with sub-Saharan Africa (SSA) and southern Asia carrying the heaviest burden (Figure 2) [9]. To achieve SDG target 3.2 (ending preventable deaths of newborns and reducing neonatal mortality to at least 12 per 1,000 live births), counting births and deaths is crucial, particularly in LMICs where neonatal mortality is highest [27], [28]. SSA alone accounted for 57% of these deaths, with a neonatal mortality rate of 27 deaths per 1,000 live births [9], [29]. The neonatal period, referring to the first four weeks after delivery, is the most vulnerable period for newborns, with a high risk of critical events happening that risk the survival of a child [10], [30]. Addressing leading causes of newborn deaths, such as premature birth, birth complications, and infections, requires QoC during pregnancy, childbirth, and the immediate neonatal period [10].

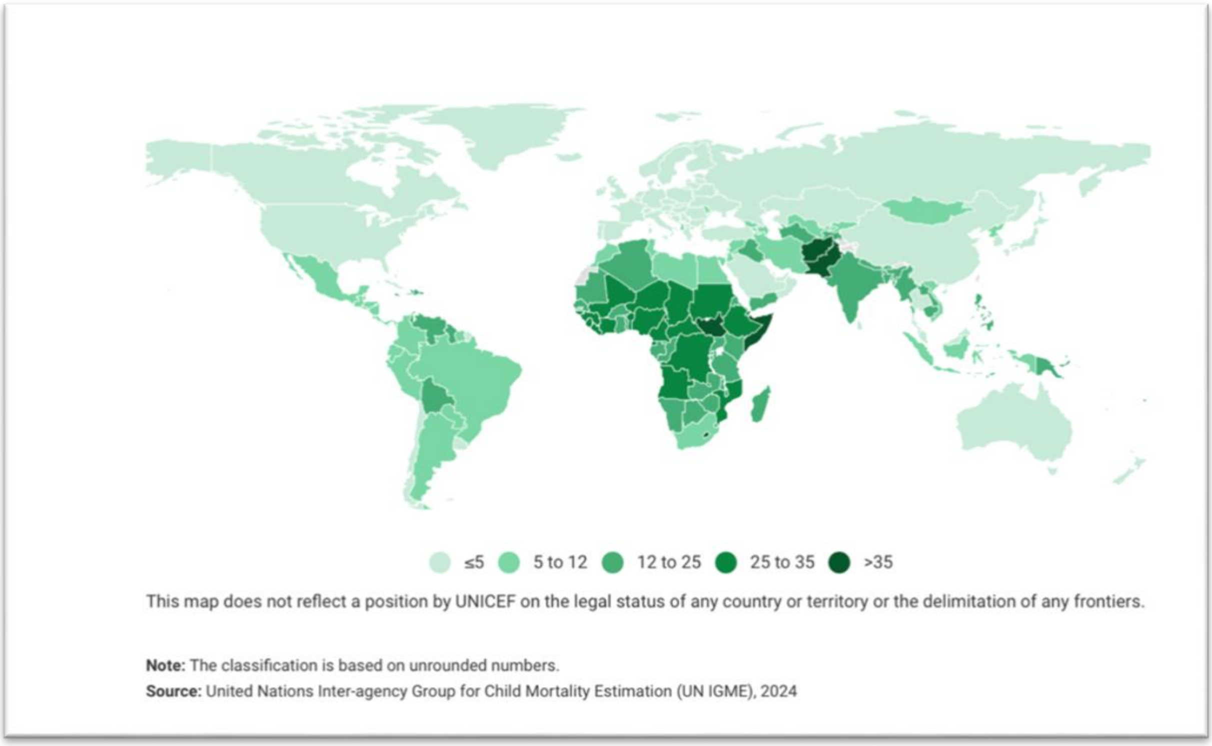


Figure 2: Mapping newborn mortality rate globally (deaths per 1,000 live births) by country, 2022, showing huge disparities across regions and countries © UN IGME, UNICEF [9].

Kenya, as part of this region, faces similar challenges with neonatal mortality due to limited access to quality care services and infrastructure. Moreover, despite the availability of high-impact interventions, many newborn deaths are still preventable [31]. The latest statistics count 21 newborn deaths per 1,000 live births in Kenya according to the 2022 Kenya Demographic and Health Survey (KDHS), a rate that varies slightly between urban and rural areas, with 21 deaths per 1,000 live births in rural areas and 26 in urban regions [32], [33] These figures reflect the ongoing challenges in preventing newborn deaths.

The Lancet Global Health Commission on High Quality Health Systems focuses on transforming health systems in LMICs by emphasising quality over access. The commission highlights that poor-quality healthcare is responsible for millions of preventable deaths each year and advocates for measuring health systems not by inputs like the number of doctors or clinics but by performance, meaning care is competent, respectful, and effective, which includes measuring health service user experience and confidence in the system [34]. To achieve high standards of QoC, health services have to be effective, safe and people-centred, as well as timely, equitable, integrated and efficient [6], [34], [35].

The WHO provides guidance and newborn care standards aiming to improve neonatal outcomes in LMICs, to ensure high-quality care during the neonatal period and reduce morbidity and mortality [36]. This includes through monitoring specific key performance indicators at the facility level. Quality indicators to evaluate newborn care in LMICs have been studied since [37], as described in the next section.

2.1.2. Key Performance Indicators for Tracking Progress

Key Performance Indicators (hereafter: KPIs) are essential metrics used to evaluate and monitor the performance and effectiveness of healthcare processes and outcomes. Process indicators help ensure quality steps are taken, while outcome indicators confirm whether those actions lead to better health results [5]. In newborn care, continuous monitoring of both types of indicators allows health systems to track progress and make quality improvements in care delivery [38].

Significant efforts have been made to develop KPIs that are suitable for different hospital settings and efficient in providing information needed to improve care. The WHO's recommendations for neonatal KPIs focus on improving maternal and newborn care by providing measurable standards for tracking progress. Several KPIs are recommended for tracking neonatal care that is required to be on a good performance level to achieve SDG 3.2, including neonatal mortality rate, skilled birth attendance, low birthweight incidence, access to essential newborn care, and QoC indicators [11], [38], [39], [40]. However, according to the Delphi study by Diego et al. (2023), there is a need for clear definitions and local guidance for facility-level newborn indicators. The evaluation of newborn indicators, based on scientific evidence, importance, and usability, led to 49 key indicators across outcome, process, and educational domains, showing that key sector actors perceived process indicators more sensitive to QoC [37]. In LMICs like Kenya, there has been progress in reducing maternal and newborn mortality. However, there remain challenges, and most newborn deaths are still preventable [9], [31]. Thus, additional indicators and efforts are emphasised, such as data collection methods and data quality, validation indicators, and the data and information systems in place [31], [41], as well as a focus on indicators linked to socioeconomic factors and equity [42], [43].

2.1.3. Health Information and Data Systems

Health systems frameworks are guiding the development and strengthening of health systems globally. The well-known health systems building blocks contain health information systems and service delivery as important blocks: service delivery needs quality, access, safety and coverage; health information systems add to well-performing systems the production, analysis, dissemination and use of timely and reliable information [44]. Health information systems are supposed to generate, process and make use of data to provide insights that are necessary to set and evaluate quality standards [44], [45], [46]. Thus, ideally, they provide the data needed to measure and manage health services and QoC.

The importance of data for achieving the SDG targets has been acknowledged since their introduction in 2015 [15], [47]. Robust health information systems that are fed with reliable data are considered critical for measuring QoC and conducting strategic planning and priority-setting [28], [48]. With the urgent need to track progress and align efforts to reduce newborn mortality and morbidity and prevent stillbirths, there has been renewed focus on the availability and quality of data, and the performance of information systems and technology to collect them [49]. However, to inform decision-making and strategies towards improving neonatal care, data also need to be translated into readable and understandable information for people at all levels of the health system, and in particular to frontline healthcare worker.

2.2. Feedback on Clinical Performance for Quality of Care

2.2.1. Audit & Feedback Interventions and Theories

Using data for Audit & Feedback (A&F) and performance interventions is a widely used strategy in this context to inform clinical practice and drive quality improvements and an approach that has been well studied. There is a myriad of research examining how to design, implement or improve existing A&F tools [2], [50], [51], and analysing whether A&F achieves the desired effects of improving QoC for patients and evaluating staff performance [24], [52], [53], [54]. These research efforts suggest that health professionals are regularly provided with summaries of their clinical performance using a range of tools (paper-based and electronic reports, feedback meetings, or displays integrated into electronic medical records (EMR)), however with varying success.

Brown and colleagues synthesised 73 feedback intervention studies to create a health care-specific theory for designing, implementing and evaluating feedback [55]. The Clinical Performance Feedback Intervention Theory (CP-FIT) provides a framework to identify key feedback characteristics and explanatory mechanisms for how it achieves influence within a cyclical process spanning: (1) goal setting, (2) data collection and analysis, (3) feedback, (4) interaction, (5) perception, (6) verification, (7) acceptance, (8) intention, (9) behaviour, (10) clinical performance improvement, and potentially (11) unintended consequences (Figure 3).

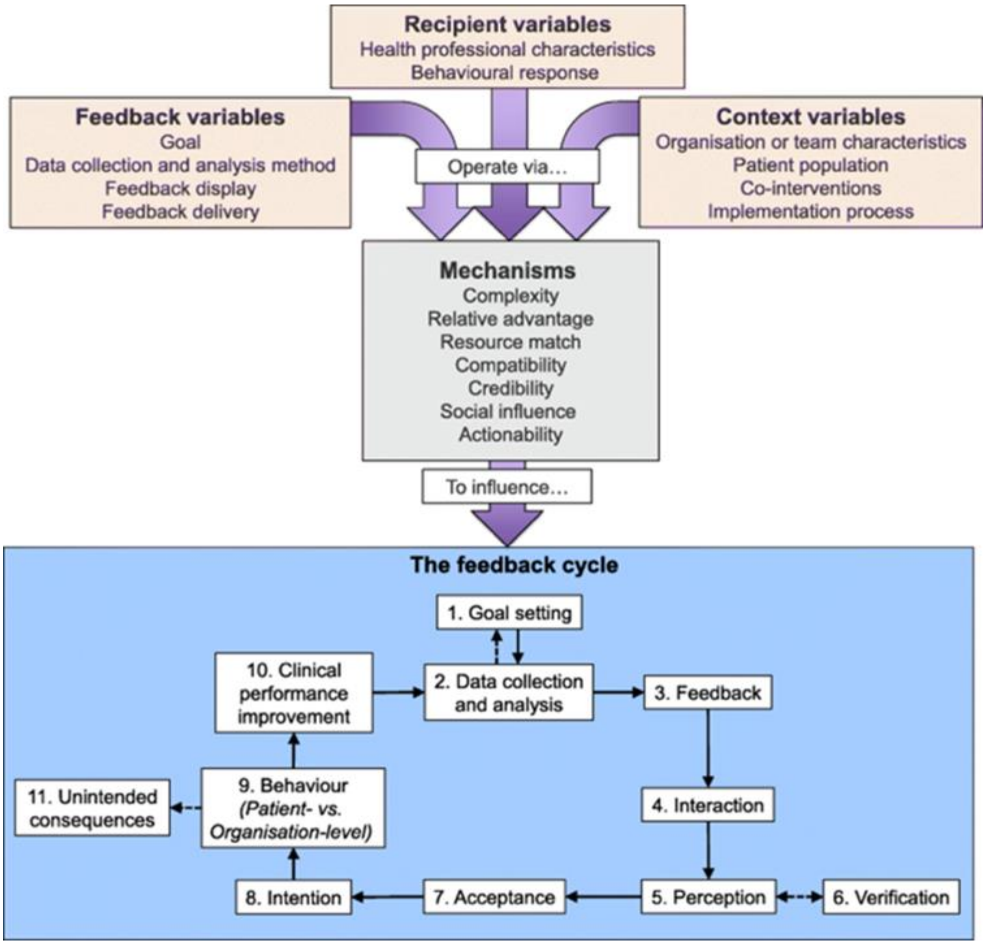


Figure 3: CP-FIT variables and mechanisms, influencing the feedback cycle. Note: This figure is reproduced from Brown et al. (2019), published in Implementation Science, and is used under the terms of the Creative Commons Attribution 4.0 International License. No changes have been made to the original [55].

Feedback interventions are less effective if any of the steps 1-10 fails, meaning if feedback display or delivery are poor (step 3), the subsequent steps (4 – 10) are influenced in a way that feedback may not work for performance improvement (step 10). As part of CP-FIT, behaviour change theories (such as, for example, Michie's Behaviour Change Wheel) play a dominant role, as taking action for performance improvement is a form of behaviour change [56]. These theories are relevant for impactful A&F interventions; however, the start is to recognise the need and intention to act. Another major influence in Control Theory, which is widely used in A&F research, emphasises goal setting and feedback loops, where performance is adjusted by comparing current behaviour and desired or set standards [57].

CP-FIT authors and developers promote their framework as a practical tool to guide or evaluate feedback interventions in the healthcare setting. To my knowledge, this framework has not been applied in the specific area of improving the quality of neonatal care in LMICs, but it could help to understand how feedback on performance through visualised information of neonatal indicators can help to improve care in LMIC hospitals. Later, Tsang, Brown and colleagues conducted a systematic review of computerised/electronic A&F systems (e-A&F) in healthcare and found that none of the included studies was from an African setting, 11% used a mixed methods design, and only 3% in paediatrics [58].

2.2.2. Data 4 Action in Newborn Care

The use of technology, such as using dashboards, can certainly facilitate transforming data into information products (e.g., visualisations) presenting results (e.g., feedback) to an audience that has neither the capacity nor the time to conduct analyses. However, as seen from CP-FIT, there are several challenges and hurdles that need to be overcome in ensuring feedback information is accurately presented and understood before clinical performance can improve. Data is only useful once processed into understanding and so presenting data and creating feedback information without taking the time to critically consider interpretation processes can lead to misunderstanding, which can cause, in a worst case scenario, no benefits but harm [16]. Thus, if we pay not enough attention to checking whether people understand feedback, action cannot result in the intended outcome or use. In a simplified 3 step process (Figure 4), I show how data translation is needed to achieve useful effects based on the CP-FIT feedback cycle. I want to highlight the importance of dashboard information presentation and interpretation as the precursors to action.

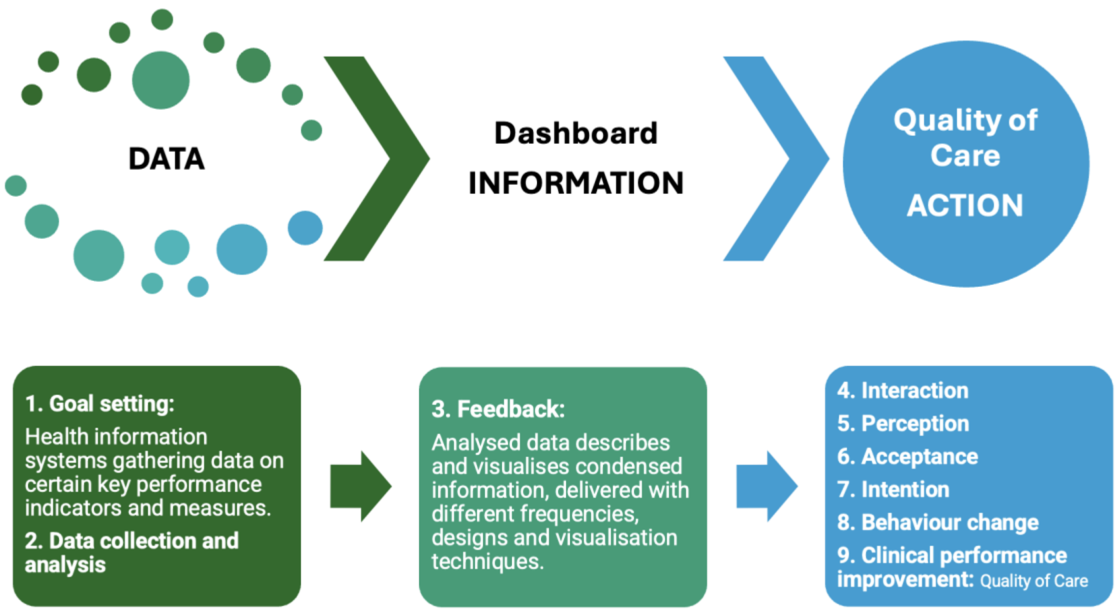


Figure 4: Data to action simplified 3-step process of data translation into useful information applying the CP-FIT feedback cycle variables to each step. The outcome is that data is used as information presentation is understood to improve outcomes. Note: Figure created by the author CH.

Health information plays a pivotal role in the performance feedback and action cycle, providing essential data to inform and improve clinical practices. Designing practical tools for using data involves setting clear KPIs that align with quality care goals. KPIs in neonatal care, such as mortality rates, infection rates, and birth complications, help to track progress and identify areas needing improvement [59]. Integrating feedback on performance to digital health tools and information systems that are currently rolled out anyway can potentially enhance practices and hence improve quality of neonatal care [60]. Therefore, investigating how data is translated into information and feedback integrated into dashboards is part of broader QoC efforts to meet global health targets.

2.3. Dashboard Design and Implementation

2.3.1. Dashboard Definition and Purpose in Healthcare

In the last decade, dashboards have been a trend in the field of information technology (IT) [25], and they become more common in healthcare to provide information to various health systems actors. With the Coronavirus (COVID-19) outbreak in 2019/20, everyone knows – or thinks to know – what a dashboard is, as many public health institutions, major media outlets, and governments started providing online dashboards to the public, tracking cases and deaths presented with world maps, bubble maps, bar charts and timelines [61]. Yet, there seem to be different understandings of what a dashboard is and what it should do. The design and implementation of health information dashboards are critical in ensuring their effectiveness in clinical and public health settings. This section explores approaches, challenges, and best practices based on recent literature on dashboard developments.

The term 'dashboard' originally described the panel on a horse-drawn carriage designed to protect passengers from mud and debris. When cars were introduced, the term evolved to describe the instrument panel displaying critical vehicle information, such as speed and fuel levels [62]. The contemporary use of the term 'dashboard' in digital interfaces represents a skeuomorphic design, replicating the function of car dashboards to present important (operating) information [63]. Thus, a dashboard today refers to a communication medium, presenting information on a computer screen monitoring 'what is going on'.

When software engineers first developed this type of technology in the 1980s, it was called an Executive Information System, with the main purpose of presenting key financial measures that "*even an executive could understand*" [25]. Thus, it facilitates the communication of data. However, the focus was on technologies rather than methodologies, on how to make data available and accessible. Afterwards, the focus shifted to collecting, correcting, storing, and accessing information in accurate, timely and useful ways. In 1992, Robert Kaplan and David Norton introduced the Balanced Scorecard, a dashboard that uses specific KPIs to drive performance strategically, compared to certain targets [64]. At the same time, dashboards have been defined loosely – for a simple reason. As long as there is room for flexible interpretation, it is easier to sell the product, in this case, the dashboard. Software company salespeople use descriptions of different features, rather than using a generic definition, e.g., it is *"able to universally connect to any XML or HTML* * *data source", "monitoring and displaying data analytics", "providing business intelligence", "generating alerts and revising actions based on KPIs",* etc. [25]. Most dashboards focus on presenting information based on tracking KPIs, which can be tailored by selecting different times (days, months, periods) and measures displayed with different types of visualisations. Often, dashboards contain multiple charts, including traffic lights and a variety of metrics [3], [25]. However, Stephen Few highlighted that a dashboard is a style of presentation, not a specific type of information or technology and consequently defined a dashboard as

> *"a visual display of the most important information needed to achieve one or more objectives, consolidated and arranged on a single screen so the information can be monitored at a glance" [25].*

No matter the dashboard type or what kind of KPIs are presented, there are common expectations of what dashboards can do, for example, to improve informed decision-making by increasing cognition and human perceptual capabilities of large, often complex databases [65] and to provide greater visibility of information and performance improvement [66]. In that sense, in healthcare settings, dashboards can be used for the purpose of feedback on clinical

* Extensive Markup Language (XML) is a markup language that provides rules to define data and is used for data representation and storage. Hypertext Markup Language (HTML) is a language used for displaying data in a web browser with a fixed set of tags, to develop the user interface [273].

performance [67], [68], [69], [70] trying to improve patient care [71] or health risk communication [72].

2.3.2. Dashboard Potential and Reasons for Failure

Dashboards have been referred to as having the potential "*to revolutionise healthcare*" by providing data-driven actionable insights [73]. They have the potential to set data into context and thus can be an effective way to bridge the gap between data and readable information. However, Dowding et al. point out that dashboards for patient care can fail their purpose to provide valuable insights due to poor design, lack of user engagement, and data quality issues [71]. Dashboards almost never fail because of their inadequate technology, but because of how they are designed and implemented, and how information is presented. Any visualisation, when misinterpreted, can fail to communicate efficiently and effectively [25]. Therefore, the potential depends on how well sound design principles and practices are followed for the dashboard design itself and for the visualisations employed [25], [74]. Since dashboards in healthcare serve as tools for visualising and interpreting complex data to support critical decision-making processes, a concise dashboard design to improve clinical outcomes is essential [70]. If well-designed, it should provide KPI visualisations to monitor quantifiable targets and actionable goals, support clinical decision-making, and deliver them in real-time [67].

However, Van Elten et al. argue that clinical dashboards can never be straightforward due to the complexity of healthcare settings. There, they are more than just data presentations; they are never static but dynamic information tracking tools that include multiple stakeholders for shaping decisions. 'Dashboarding', as the authors describe it, is a continuous process, and adjustments happen organically based on evolving needs of projects, and there can be tensions between creating one that can be used universally and one that is tailored to specific goals [73].

2.3.3. Design Principles for Hospital Dashboards

There are several, well-studied design principles that inform a 'good' dashboard design. Design concepts, from ancient craftmaking to modern product design, have evolved to combine functionality and aesthetics in digital solutions. Design as a discipline has gained momentum in the 20th century, when simplicity and functionality helped to better-selling products. Henry Dreyfuss highlighted the importance of designing for people, and later Don Norman advanced this approach to user-centred design principles [66],[76]. Integrating user experiences and iterative prototyping have become important parts of the design process since, in particular for technologies and software. Thus, A&F dashboards in healthcare may benefit from applying these design principles too.

2.3.3.a. Human Computer Interaction & User Centred Design

Human Computer Interaction (HCI) is a multidisciplinary field focussing on how users interact with computers or smartphones where information is taken from, thus includes design aspects. HCI aims to improve the computer technologies' usability and user experience by understanding the cognitive, social, and organisational aspects of the users' interaction with them [77], [78]. HCI tries to recognise and accommodate different user groups, such as developers and end-users, in their interactions with computer technology (Figure 5).

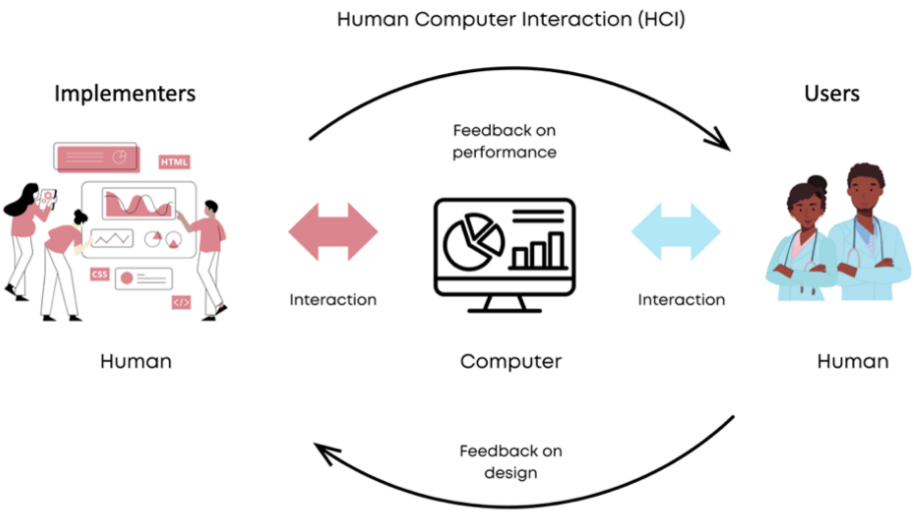


Figure 5: Human Computer Interaction (HCI) to visualise that developers (implementers) and users interact with dashboards, although they may interact differently. Note: Figure created by CH.

HCI is an extensive field that intersects with and has laid the foundation for design disciplines. User involvement and assessing design for user needs is something that various design principles aim towards to [79]. Human-Centred Design (HCD) or User-Centred Design (UCD) – both terms used interchangeably – is an approach within HCI that looks at user needs, preferences, and limitations at every stage of the design process by including the end-user in the design process [26], [76], [79]. UCD puts the user at the centre and uses an iterative approach to make sure real-word needs are met, the tool, product or service is useful and easy to use and tries to ensure that technology is adapted to the user's environment rather than forcing users to adapt to the technology [26], [80].

ISO Standards define HCD/UCD as

> *"an approach to system design and development that aims to make interactive systems more usable by focussing on the use of the system, applying human factors, ergonomics and usability knowledge and techniques"* [80].

Don Norman goes further and defines HCD/UCD as

> *"a collaborative process, putting people at its centre and designing around their needs, capabilities and behaviours"* [76].

Both, HCI and UCD, incorporate to enhance user experience (UX), user interfaces (UI), usability, and satisfaction, aiming to create effective and user-friendly (digital) products [81].

2.3.3.b. User Experience, User Interface & Usability

While often used interchangeably, they each have distinct roles and best practices in the design process [82], [83], [84]. UX is about the user journey, including usability, accessibility, and overall satisfaction when using the product, service or system [83]. UI refers to the design framework focusing on the visual elements and layout, as well as the looks and styles desirable by the user [84], [85]. However, UX and UI are interdependent (Figure 6) to create user-friendly outputs [82].

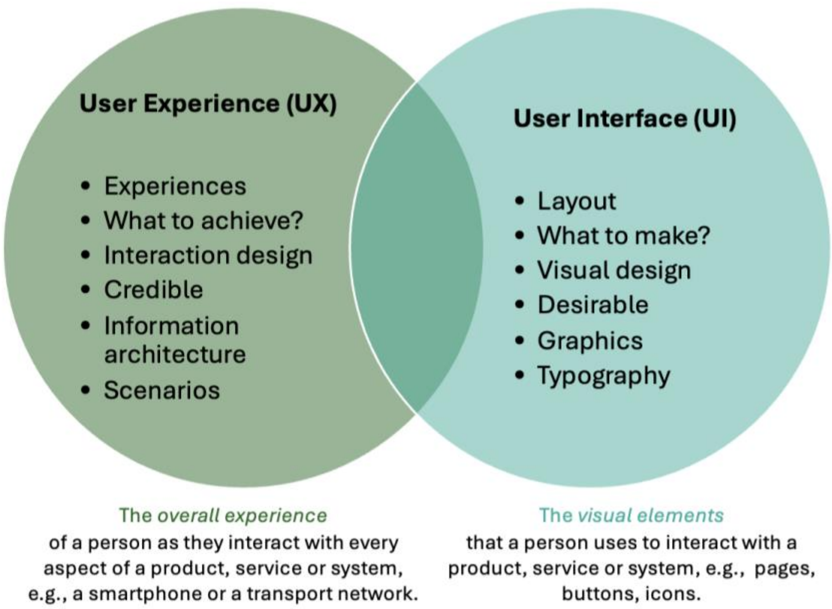


Figure 6: User Experience (UX) and User Interface (UI), closely linked but different roles: UX is about the experience and satisfaction of the user, while UI focuses specifically on the visual design, aesthetics, and interactive elements that contribute to that experience [83], [84]. Note: Figure created by CH.

Usability, a part of UX, adds a deeper level, measuring how effectively, efficiently, and satisfactorily a user can interact with a product, often evaluated through user testing or expert reviews (Figure 7). However, both theories guide the development of easy-to-use and useful products, services or systems, reducing the risk of errors [86].

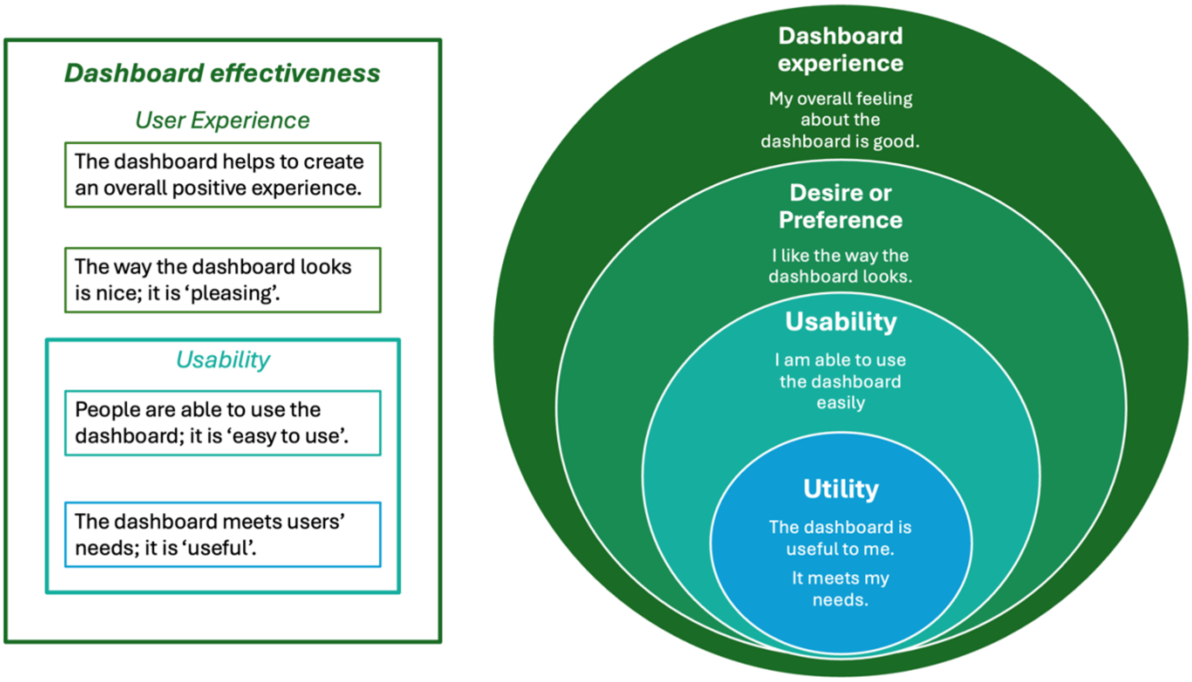


Figure 7: Dashboard user experience containing several layers to understand usefulness. Note: Figure created by the author CH, inspired by visualisation from 'User Experience Conference Amsterdam 2008', NN Group' [87].

Catering for user experience is more than giving people what they say they need [88]. For dashboard users to have an overall positive experience when using a dashboard, this can include user preferences, the functionality, and the utility of the dashboard, meaning it pleases the user and is 'easy to use' and 'useful'. Over time, UCD and UX approaches have evolved beyond usability to consider the overall quality of interaction between users and a product, system or service due to the importance of users' emotional experience. The rise of empathy in design, particularly through empathic design and design thinking, led to an increased focus on understanding users' needs, emotions, and contexts of use, and to adopt participatory, co-design methods, meaning users are directly involved in the design from the start [86], [89].

Techniques like sketching (individual or aggregated) user empathy maps capture what users say, think, do and feel when interacting with the solution, placing the user at the centre (see Figure 8), which allows designers to understand the users' world [89], [90], [91].

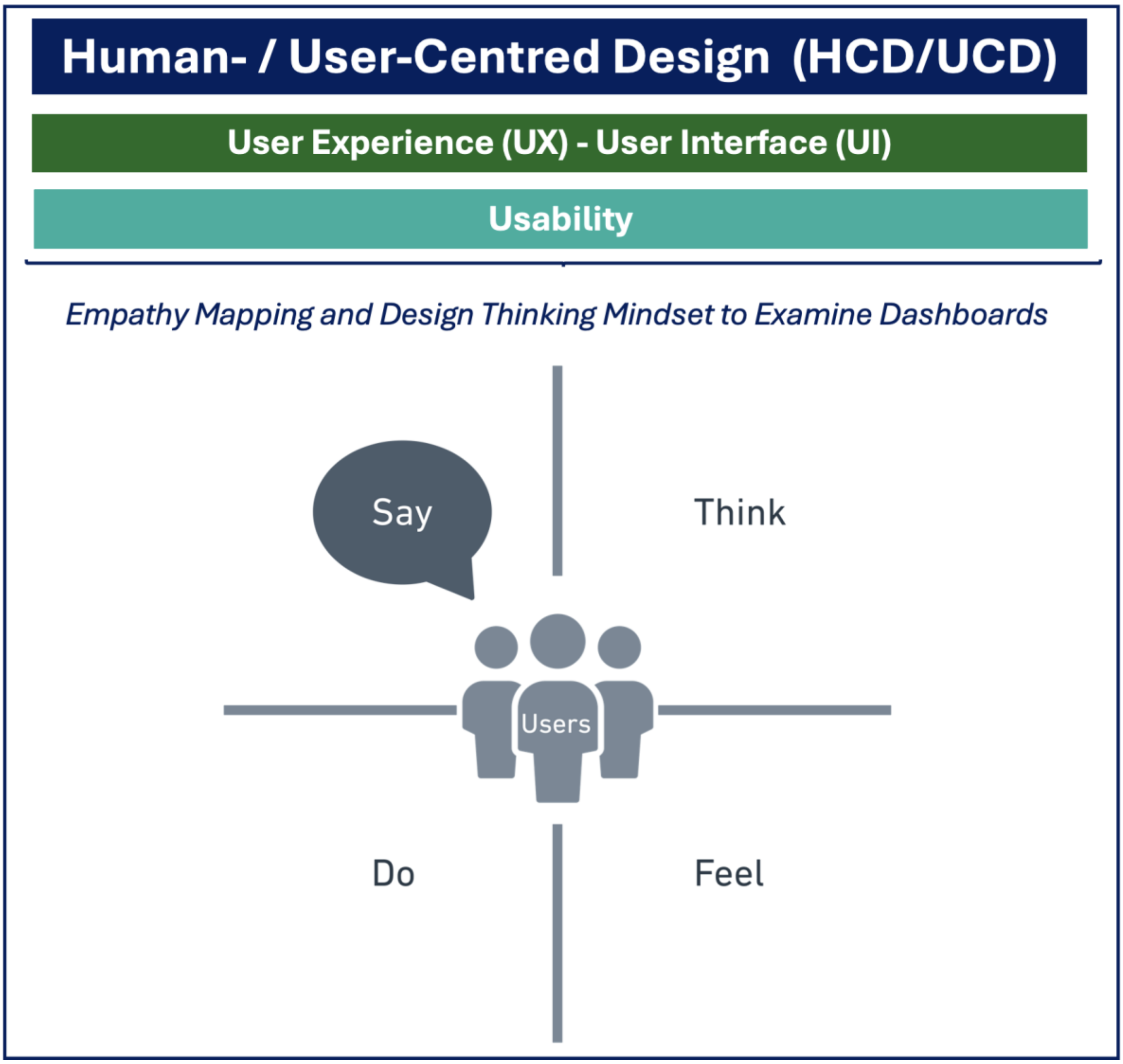


Figure 8: Design principles, empathy mapping and design thinking mindset for exploring dashboard user experience. Note: Figure created by CH, inspired by the User Empathy Map, NN Group [89].

However, it is also known to be a challenging task to balance designers' and users' knowledge (design for and by users), and direct involvement is not always feasible and not always necessary. Balancing knowledge requires creating equal partnerships through co-creation, combining user-centred and participatory design, to facilitate balanced interactions (e.g., through effective recruitment, reflection, and managing team diversity) [26], [76]. Don Norman, like all other design specialists, furthermore stress that design is never linear, but a continuous and iterative process. Thus, it is essential for improving usability to test, refine, and incorporate user feedback. Design is never perfect initially, and iterative testing with users ensures the product better aligns with their needs and reduces errors [76], [86]. Figure 9 shows how different steps in the design process lead to improvement through iterations.

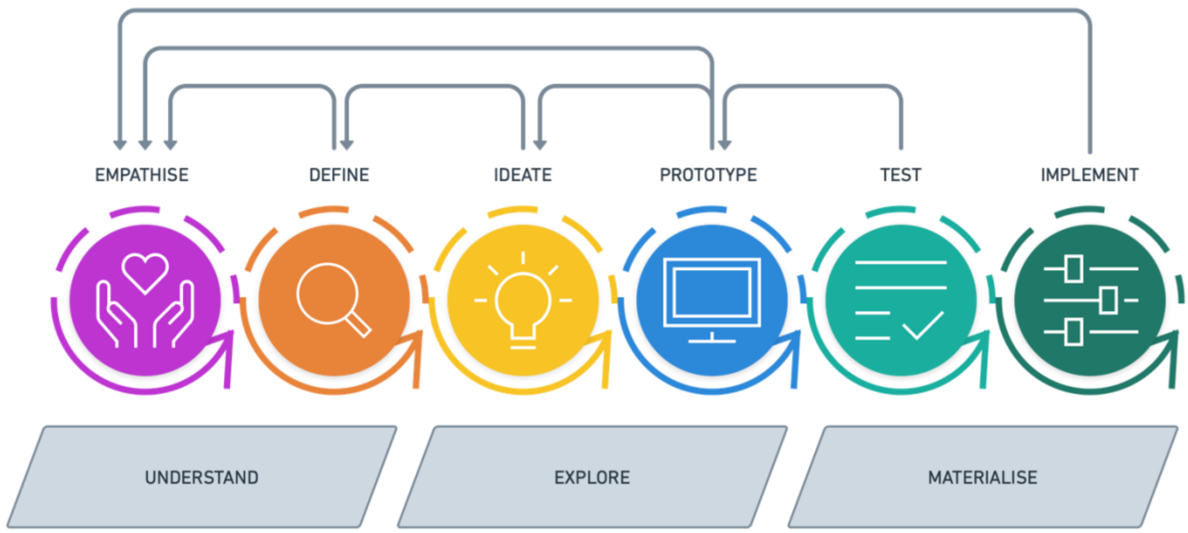
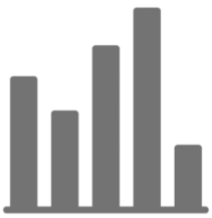
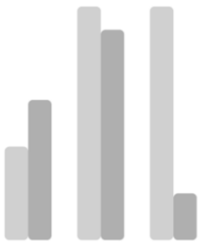
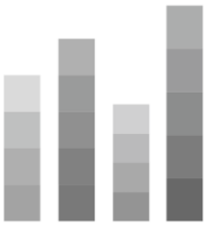


Figure 9: Design process flow involving six steps: (1) empathising with users to understand their needs, (2) defining the core problem, (3) ideating possible solutions, (4) prototyping to explore these ideas, (5) testing to refine the solution based on user feedback, and (6) implementing and evaluating the solution. Visual created by CH, inspired by and adopted from the Design Thinking model by the Nielsen Norman Group, with changes in colours and style [102].

2.3.3.c. Effective Data Visualisations & Graphical Perception

A visualisation is usually a visual representation of information. While *data visualisation* is used as the umbrella term for all types of visual representation, the subset *information visualisation* (and *scientific visualisation*) refers to a visualisation type with a particular purpose, not only to create a picture but also to help us to think and understand [93]. Accordingly, Card, Mackinlay and Shneiderman extended the definition of information visualisation as "*computer-supported,* *interactive, visual representation of abstract data to amplify cognition*" [85], [94]. Charts, graphs or maps - i.e. a pictorial, graphical representation or diagram that represents a set of numerical or qualitative data or values – but also tables and even text, if sufficiently organised, can serve as a form of information visualisation [87],[74]. Measures can be done as ranking, deviation, correlation, distribution, change over time, spatial, part-to-whole or flow [96], [97], [98]. However, with more and more data and digital tools to process them, it can be challenging to decide which visual presentation works best to amplify cognition. Table 1 presents an overview of ubiquitous graphical presentations, describing the use and flaws (often not recognised by those using graphical tools).

Table 1: Overview of commonly used Information visualisation types, inspired by Knaflic [74], Smith [99], and Few [100]

Graph	**Name**	**Description/ Use and flaws**	**Application**
	**(Column) Bar Chart**	A bar chart displays categorical data using rectangular bars whose heights or lengths correspond to the values they represent. It is the typical method for comparing data. When arranged in order, standard bar charts make it easier to see the ranks of values. Column bar charts are effective for illustrating changes over time, especially when depicting only one series of data at a time. All bar charts can be created both vertical and horizontal.	Ranking, Distribution, Change over time, Magnitude
	**Paired (column) Bar Chart**	Like a bar chart above, a common method for comparing is presenting data as paired or grouped bar charts, accommodating multiple variables/series. Reading can become difficult when dealing with more than 2 bars. Bar charts and paired or grouped bar charts are commonly used and easy to understand. However, it's important for these charts to show a zero baseline, or else it could lead to inaccurate visual comparisons.	Ranking, Distribution, Change over time, Magnitude
	**Stacked Bar Chart**	Bar graphs that use stacked segments within their bars are used for displaying the division of a larger category into smaller categories and the impact of each part on the total amount. They are suitable for presenting survey results that involve sentiment, such as disagree, neutral, or agree. As the number of segments in each bar increases, one major downside is that they become harder to interpret. Comparing each segment to one another can be challenging since they are not aligned on a common baseline, and it requires more cognitive effort.	Deviation

	**Diverging Bar Chart**	Like the bar charts above, a diverging bar chart can accommodate both negative and positive magnitude values. The primary purpose of a diverging bar chart is to compare data with a middle point or a baseline. It can be effective for illustrating the range of negative and positive values, such as displaying survey results that include sentiments from strongly disagree to strongly agree. However, comparing the values of the categories in the middle can be challenging and requires more cognitive effort.	Deviation
	**Line Chart**	A single series of data, two series, or multiple series can be displayed on a line (or multiple line) graph. It is often used for presenting continuous data and illustrating a changing time series. The line connecting the points implies a relationship, which may not be suitable for categorical data. If the data is irregular, it's important to use markers to represent the data points.	Change over time
	**Column Bar and Timeline Chart**	Plotting various data against a common x-axis can be a powerful way to present data. It can illustrate the correlation over time between a quantity (columns) and a rate (line). However, using two distinct y-axes increases the likelihood of misinterpretation and it requires more time for readers to discern which data corresponds to which axis.	Correlation, Change over time
	**Scatterplot Chart**	A scatterplot is a visual representation that uses dots to depict values for (usually two) numeric variables. It is effective for illustrating the connections between two continuous variables, with each having its own axis. While frequently employed in scientific contexts, some individuals unfamiliar with this graph type may find it complex.	Correlation
	**Pie Chart**	Pie charts are commonly used to represent part-to-whole data, but they can make it challenging to make accurate comparisons between segments. They are usually used for presenting data categorised into nominal or ordinal groups, typically showing percentage or proportional data, with each category's percentage often displayed next to its corresponding pie slice.	Part-to-whole relation
	**Donut Chart**	Like a pie chart, but the centre of a donut chart provides space for additional information, such as the total. Both, pie and donut are not the most effective method for conveying information because it is more challenging to compare segments. Due to the absence of a scale, it is challenging to accurately interpret values. The basic premise is that they are not effective for communicating data. They also occupy more space.	Part-to-whole relation

There is a large body of literature from experts in cognitive psychology and data visualisation on how a user's visual system decodes the information in a graph and what graphical perception tasks are performed with greater accuracy than others [101], [102], [103], [104].

The so-called *Seven Gestalt Principles of Human Perception* describe how humans group similar elements, recognise patterns and simplify complex images when they perceive objects, which applies on visualisations. Designers use the principles to organise content on websites and other interfaces so that content is not only aesthetically pleasing but also easy to understand [101], [105]. Functions and forms, as well as colours, size and position on a page can attract or distract attention. To keep the audience's attention and to minimise misinterpretation, 'workload' needs to be kept low, meaning to remove anything that takes up brain power unnecessarily. A good way to keep attention is to create a hierarchy of information, meaning the user will be guided through different parts of information in a specific order according to its importance or logic [74].

In understanding how humans process and interact with information, various graphical perception and cognitive theories describe how to design and visualise based on natural cognitive abilities and cognitive psychology. For example, George A. Miller's theory on working memory suggests that humans can hold 7 ± 2 objects in short-term memory, but chunking information into larger units enhances retention, while K. Anders Ericsson's research found that there are expert learning methods and mental representations for superior performance [100]. Cognitive Load Theory suggests minimising mental effort in system design to reduce cognitive overload can improve usability [86], [104].

How useful a visualisation is, depends on how much **context** is considered: Who is presenting to whom, who is the audience of presented information, in which setting do people use this information, and what are the actual information needs? These are questions that have to be answered for visualising data for information to become successful communication [74]. Needs and interests of different groups can be mutually exclusive. Any group, community or individual has information needs and uses information in relation to a specific purpose and within a specific real-world context (e.g. employment, social or cultural activities, studies etc.). However, the question is whether people always recognise an information need when they have one. When asking people about their information need, it is also important to carefully differentiate between *“this is what I was told to be my information need versus this is my actual information need in reality”* [106].

Finally, the concept of **primary and secondary users** is central in UCD and equally relevant for creating effective data visualisations. The concept refers to distinguishing between those who are the intended users that use the tool more regularly (primary users) and those who might as well interact with it in addition somehow, however less frequently (secondary users). This helps to reduce the risk of limiting usability and accessibility to only one type of user and ensures that the interface is functional and inclusive. [25], [26].

2.3.4. Implementing Digital Feedback Interventions

For A&F interventions and dashboards, just as for making use of data through visuals, the context in which they are implemented and used is a key aspect for (re-)designing them according to users’ needs. Most of the studies and theories highlighted above, where developed and applied in a HIC setting, where the infrastructure and resources differ significantly from the LMIC settings. Although these insights can help to inform what works, the hospital setting can vary a lot between countries, and even within countries between regions and levels. In the specific context of Kenyan public hospitals, implementing dashboards requires consideration of the organisational structure and data workflows, which are often complex, and mostly still paper-based [107],[15].

Even in well-resourced settings, challenges arise related to integrating dashboards into existing systems, ensuring data quality, and maintaining user engagement over time. A successful implementation requires continuous assessment and adaptation, especially to meet user needs and contextual demands [108], [109]. In LMICs, the implementation success can further be impacted by resource limitations, infrastructure gaps, and varying levels of technological adoption. Hospitals may face challenges such as unreliable electricity, limited internet access, and staff shortages. Thus, implementing hospital dashboards or e-A&F for quality improvement in settings where digital health interventions can be complex to implement and where further evidence is needed may involve additional difficulties, particularly around contextual adaptation, scalability, and sustainability [108], [110], [111]. Many digital health interventions, in fact, have difficulties in scaling because they are not adequately tailored to the specific social, political, economic, and infrastructure realities of the hospitals and regions in which they are deployed, which affects the long-term viability and widespread adoption [112], [113]. In addition, for sustainability, dashboards just as other digital health interventions require continuous support, maintenance, and engagement with healthcare workers and leaders [111]. While dashboard challenges are not exclusive to LMICs, as HICs also face difficulties in dashboard adoption, there is a noticeable gap in the literature on African countries hospital dashboards, especially those focused on newborn care.

2.4. Thesis Contributions to Existing Evidence

Here, I present how my thesis contributes to the gap in the literature and the practical outputs that are needed, the scope and aim of my research and what research questions my research will answer.

2.4.1. Academic Contributions and Practical Impact

My research intends to address a gap in the literature by taking a specific focus on data sense-making (how to use data to generate understanding). The results will then contribute to clinical practice by providing guidance on how data needs to be translated using digital visual presentation for specific audiences in the context of neonatal care improvement efforts in LMICs so that they can make best use of information. In this way, I will address an often-overlooked challenge in the use of feedback to improve performance [114], [115].

There is a large body of literature on feedback interventions and dashboards in general, but very little with a specific focus on dashboards used for improving performance related to neonatal care in LMICs. And while there is a myriad of literature on how dashboards should be designed and implemented in general, there is no consensus on a general definition of what constitutes a dashboard, including its technical systems, specific functions, and expected outcomes. My research addresses this gap by providing detailed guidance and key considerations for designing and implementing an A&F tool/ dashboard developed to support actual change and quality improvement in neonatal care. By exploring specific tool design and implementation aspects, this study contributes both to theoretical frameworks in A&F and practical applications for dashboard interventions. The research draws on experiences and lessons learned from two dashboard interventions focused on the health system in Kenya, offering valuable insights into the challenges and successes of integrating digital dashboards in LMIC healthcare settings. These contributions not only advance the academic discourse but also provide actionable strategies for healthcare practitioners and policymakers aiming to enhance care quality through effective use of digital health tools aimed at improving performance.

2.4.2. Study Scope

This research benefits from collaborations with two initiatives: the KEMRI Wellcome Trust Research Programme's CIN (for: Clinical Information Network) and the NEST360 (for: Newborn Essential Solutions and Technologies) programme. CIN brings together a set of Kenyan hospitals and medical and nursing leaders from newborn units to form a network with an intervention approach to improving service delivery [116]. NEST360 (shortened hereafter to 'NEST') is a multi-institutional partnership programme deployed in Kenya, Malawi, Nigeria and Tanzania committed to reduce neonatal mortality. Both CIN and NEST are trying to use data for action in newborn care providing performance feedback on key process and outcome indicators to their partner hospitals. These initiatives have developed dashboards to deliver feedback on analysed data to hospital staff [117], [118]. Focussing on how information is presented, my study uses these dashboards to investigate how they translate data into actionable insights and support hospital staff in their efforts to improve quality of newborn care (Figure 10). More details about the two initiatives as collaborators will be presented in Chapter III.

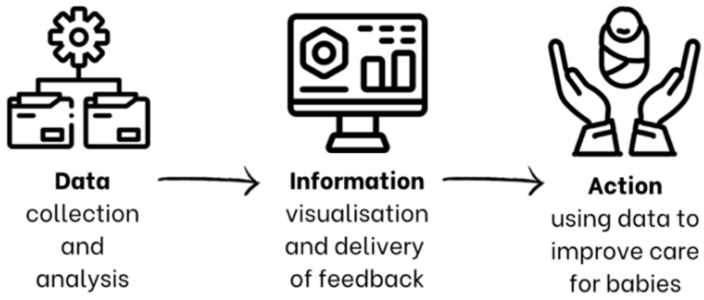


Figure 10: Simplified 3- step process of data translation into useful information so that action can happen as in data in form of graphs is translated to usable information so that action is data-driven. Note: Figure created by CH.

2.4.3. Aim and Research Questions

The overarching research aim is to examine how more available and larger amounts of health data may enable to use data for performance feedback at various levels of the health system and engage multiple actors in efforts to drive improvements in neonatal hospital care in LMICs, particularly in Kenya. This contains an academic and a practical goal to understand how to improve effectiveness of data-driven feedback dashboards, and how these can be implemented and adopted at a wider scale. Using CIN and NEST dashboards implemented in Kenya as the basis for the empirical work, the following research questions will be addressed:

Question 1	**How have dashboards been used in neonatal care and what can be learned about effective dashboard designs?**
Question 2	**How well do key health sector actors understand and interpret feedback from dashboards?**
Question 3	**How do health sector actors use information provided by specific feedback tools in their daily work, and what are the implications for daily practice in newborn care?**

Question 1 focusses on identifying evidence and best practices from published research, assessing how dashboards are developed to provide newborn care relevant information. Question 2 will then focus on information presentations of CIN and NEST dashboards to investigate how well actors from various target groups understand and interpret different feedback displays and visuals. This includes to ask what their efforts and needs are to understand feedback, and what their preferences are for different indicators or types of feedback, and tools. Finally, question 3 will examine the implications of different feedback visualisations and dashboards in practice, how it motivates users or helps to know what or how to act on it. Answering these questions will help to understand how to optimise feedback strategies to help improve performance in networks for neonatal care that can be applied more widely across health systems, especially as these become more advanced in LMICs.

Chapter III. Overall Approach and Methodology

This Chapter outlines the overall design and methodology of the proposed research, including the philosophical perspective, study framework, selected methods, the study setting, and collaborators. It concludes with positionality, reflexivity and ethical considerations.

3.1. Research Design and Questions

Located in the field of health policy and systems research, the study design drew on the rapidly emerging body of implementation research [119], [120], and more specifically on A&F theories and design principles as described in Chapter II. This study's goal is to provide insights how to optimise feedback design and implementation strategies for newborn dashboards, to improve performance in professional networks for neonatal care. Figure 11 provides an overview of the selected study methods to answer the research questions. In this chapter, I describe the scoping review and mixed methods design, and how research activities were informed by CP-FIT and design principles.

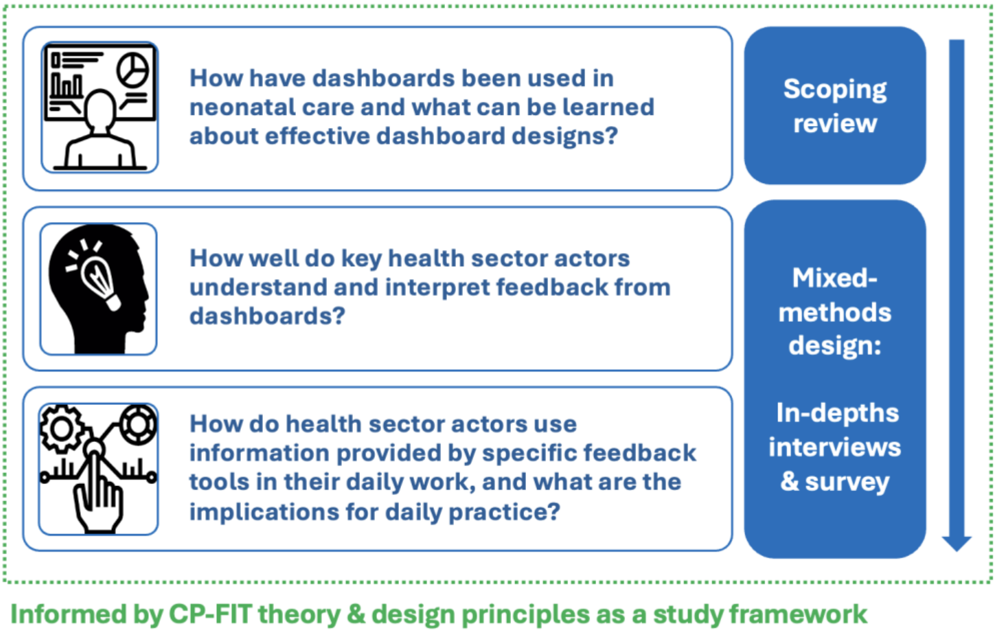


Figure 11: Research questions and selected methods to answer them.

3.2. Philosophical Perspective

3.2.1. Research Paradigm

The philosophical perspective guiding this research is rooted in pragmatism, which applies practical solutions to problems based on real-world experiences. I seek to understand reality through practical interactions with the data and technology, aiming for solutions that address real-world challenges in the context of healthcare feedback systems and (digital) health interventions [121]. Pragmatism views reality as something that is not fixed but can be shaped by human actions and decisions [122], [123]. Epistemologically, knowledge is gained through practical engagement and observation, focusing on what works in a given context. Pragmatism is often described as a philosophical foundation for mixed methods research as it supports the use of different research methods and a continuous cycle of inductive, deductive, and abductive reasoning. This cycle can lead to more comprehensive, culturally relevant and socially useful knowledge [122], [123].

I integrated complexity theory into my study to account for the unpredictable and interconnected nature of health systems. Health service interventions operate within complex adaptive systems, where numerous agents (patients, providers, implementers, technologies and tools) interact in non-linear ways. Complexity theory captures measurable outcomes as well as emergent behaviours and contextual factors that influence the success or failure of these interventions [124]. By combining complexity and pragmatism in my study, I could remain responsive to changing conditions and grounded in real-world applications. After all, the aim was to ensure that findings are actionable and relevant for improving quality of hospital care for newborns in Kenya and potentially other LMIC settings.

3.2.2. Theoretical and Conceptual Perspective

As described in Chapter II, CP-FIT by Brown et al. (2019) is an evidence-based theoretical framework designed to improve the effectiveness of feedback interventions in healthcare settings based on pre-existing feedback interventions and behaviour change theories. In my study, CP-FIT served as the core theoretical framework because it provides an organised, evidence-backed model based on established theories and outlines mechanisms and pathways that explain how and why feedback interventions lead to behaviour change in healthcare professionals [55]. Furthermore, as CP-FIT is based on multiple analyses of interventions, mechanisms and theories, it offers flexibility similar to a conceptual framework and can be applied in various healthcare contexts.

A conceptual framework is based on a collection of concepts and ideas that guide the research's operationalisation and the relationships between variables, which offers flexibility to adapt as the research progresses. In contrast, a theoretical framework connects the research to existing theories, which provides a stable foundation grounded in established knowledge [125], [126]. A well-aligned study systematically connects the research questions, data collection, and analysis to meet rigorous research standards [125]. Integrating both conceptual and theoretical frameworks, as recommended by Ravitch & Riggan (2017) and Marshall & Rossman (2016), can strengthen the study by combining flexibility and theoretical depth and by bridging the gap between theoretical insights and practical research, and thus enhance the overall rigour. This mixed approach is especially useful in complex, interdisciplinary research and offers a comprehensive foundation for deeper analysis and understanding [126], [127].

Using CP-FIT flexibly helped me design my study based on a well-developed theoretical structure and let me add depth by including insights from other literature bodies and theories to address areas that CP-FIT does not extensively cover. For example, I drew on literature on visual information science, human-computer interaction and human user-centred design, which provide relevant concepts for designing and implementing healthcare interventions and technologies as seen in Chapter II. By incorporating a broader body of literature, I minimised limitations that can arise from relying on a single framework. Remaining open to new ideas from relevant fields gave my thesis more rigour, which also aligns well with the pragmatic, real-world approach needed for actionable and relevant findings in complex settings like healthcare, and specifically quality improvement interventions [124], [127].

3.2.3. Study Framework

Using CP-FIT, the thesis investigates the practical issues of dashboards for improving care for newborns in Kenya. I used the theory to go through the feedback cycle, investigating how goal setting and different feedback displays affect interaction, perception, acceptance, intention and behaviour, examining what dashboard users understand from information visualisations, how they interpret them and what their needs or preferences are for taking action. I then explored what the influences are for dashboards to be effective feedback interventions to improve QoC performance. This helps to understand the actual use of data and whether the format in which information is presented influences actors' use of feedback at different levels of the health system, and how this in turn may affect the whole feedback cycle (Figure 12).

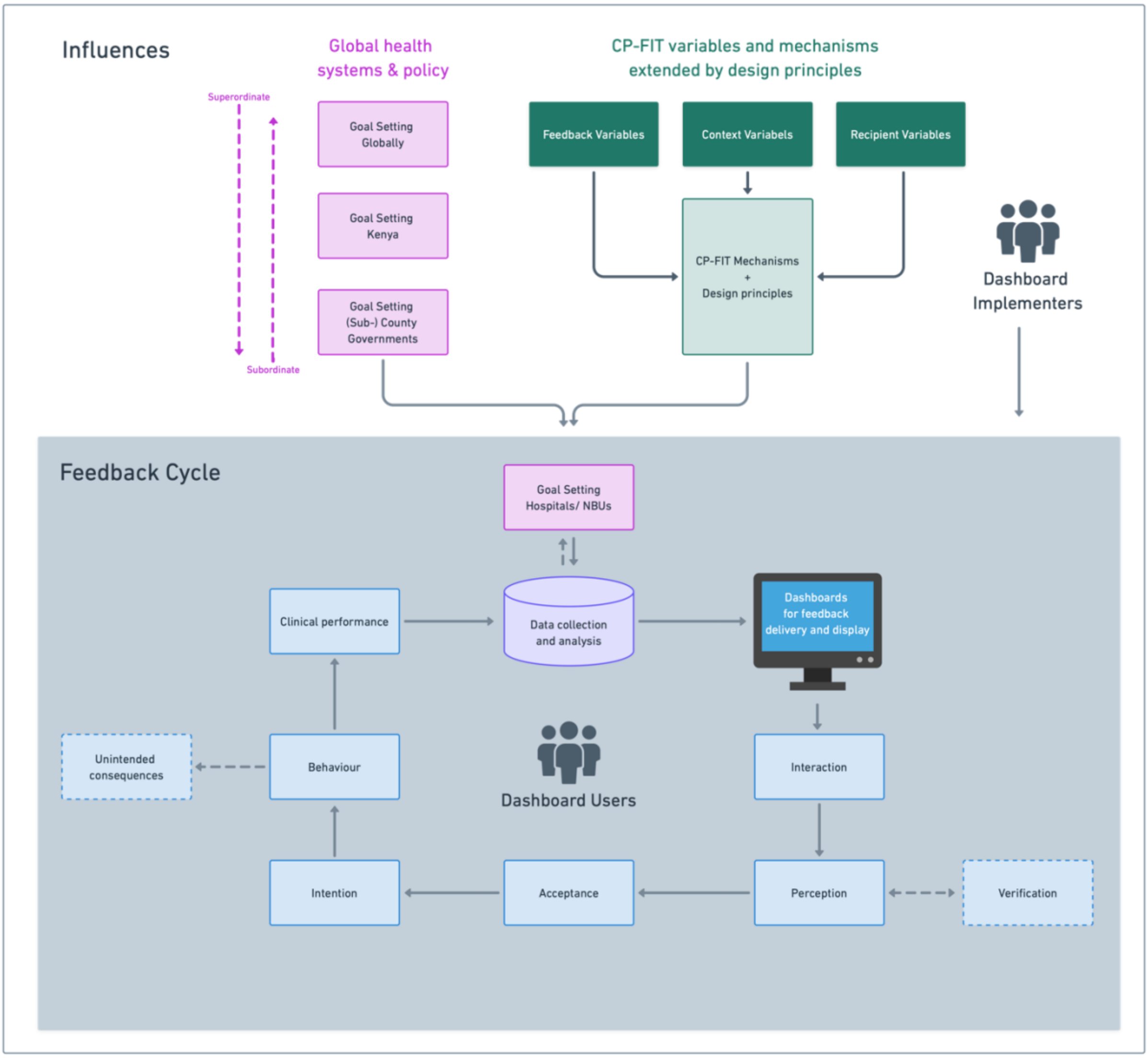


Figure 12: Adjusted CP-FIT variables, mechanisms and feedback cycle, extended by a global health systems and policy perspective, a dashboard user and implementer perspective, and other design principles and theories adding to CP-FITs mechanism that shape performance dashboard use in LMICs. Note: Visual created by CH, reproduced and inspired by CP-FIT figure 3 from Brown et al. (2019), published in Implementation Science, and is used under the terms of the Creative Commons Attribution 4.0 International License. No changes but additions have been made to the original [55].

CP-FIT is informed by 73 feedback interventions and 30 pre-existing behaviour change theories, such as context and implementation theories, feedback theories, general behaviour change theories, motivation theories, psychological theories, sociological theories and technology theories. Although usability is mentioned as part of the technology theories, I extended CP-FIT using knowledge and concepts from the wider design literature to examine the dashboard experience comprehensively from a user perspective. These included human-computer interaction and human-/user-centred design principles, with a specific focus on the different layers of user experience, user interface, usability and user needs as seen from Chapter II, Section 2.3.3.

Brown et al. uses the terminology “feedback recipients” [55], however, I will refer to the feedback recipients in this thesis as dashboard **users** and the feedback providers as dashboard **implementers.**

3.3. Selection of Methods

The methods were carefully selected based on a narrative review and my observations in Kenya prior to this DPhil research. I considered alternative methods and their limitations, focusing on what methods were most suitable to answer my specific research questions and the study context. The research questions were answered using the different methods outlined in Table 2 below.

Table 2: Research questions related objectives, methods and target groups.

Research question (RQ)	Related objectives	Methods	Target groups/ participants
RQ 1: How have dashboards been used in neonatal care and what can be learned about effective dashboard designs?	**Mapping exercise** of existing evidence on newborn dashboard designs: to identify research on dashboards for tracking neonatal indicators across various settings to learn which factors, methods, design guidelines, and KPIs were effective in improving newborn care.	**Scoping review** to comprehensively map a broad body of literature and include a wide range of study designs, as a first step to inform the empirical part of this study and to guide future evidence-based dashboard development.	**Eligible studies** include any information management tool that visually tracks, analyses or displays newborn key performance indicators (KPI), metrics and key data points to monitor newborn health or specific newborn related processes with electronic tools.
RQ 2: How well do key health sector actors understand and interpret feedback from dashboards?	**Empirical research** to examine how key health sector actors understand and interpret visual feedback from dashboards, and to identify their preferences and needs to effectively act on the information provided.	**Mixed methods research:** Quantitative survey to assess dashboard content and qualitative in-depth interviews with dashboard users and implementer, including a dashboard talk-through with users.	**Participants** are recruited from Kenyan health sector, healthcare workers involved in newborn care at Kenyan public hospitals, CIN and NEST users and implementers.
RQ 3: How do health sector actors use information provided by specific feedback tools in their daily work, and what are the implications for daily practice in newborn care?	**Empirical research** to explore how health sector actors then use information from specific feedback tools in their daily work and to assess the implications of this usage for practices and implications across different levels of the health system.		

Detailed explanations and justifications for the specific methods used will be provided within each findings chapter: the scoping review (Chapter IV), survey (Chapter V), and interviews (Chapter VI). In this chapter, I will outline the mixed methods design for the empirical work.

3.3.1. Mixed Methods Design

A mixed methods approach is a research methodology in its own right. It combines both qualitative and quantitative research methods in a single study to combine the strengths of each. This approach allows research questions to be explored from multiple angles, providing richer and more actionable or complementary insights, particularly valuable in complex settings like healthcare [128]. Mixing two methods can be more powerful than a single method in allowing insights into context and understanding the 'how' and 'why' of phenomena through incorporating the narratives of individuals and groups [122], [125].

In my study, the rationale was to answer research questions 2 and 3 using both quantitative (closed-ended) and qualitative (open-ended) data. I included two newborn dashboard initiatives providing feedback on neonatal indicators to different target groups (clinicians, hospital managers, programme managers and governmental actors). The selected initiatives are described in greater detail below under setting and collaborators. The aim was to learn from users and implementers about how these tools support the use of data to improve neonatal hospital care, specifically regarding:

- the 'ease of use' and 'usefulness' of these tools,
- user needs and preferences for design and delivery of information, and
- how feedback tools motivate or guide to improve QoC.

The findings were intended to help optimise (electronic) feedback tools for neonatal care in Kenya and potentially other LMICs.

3.3.2. Combining Selected Methods

For the quantitative component, I selected an online survey and for the qualitative component, I conducted in-depth interviews.

The survey assessed how users understand and interpret visual feedback from different graph types currently used in the two dashboard feedback tools. The survey was used to test visual decoding and understanding of the feedback tools by asking users to interpret information presented with different graph types retrieved from both tools. Furthermore, participants were asked to indicate their personal preferences and perceptions (e.g. how useful and easy to use the tool graph is). The intention was to identify how the translation of data to information presentation of performance indicators can be improved, which elements would benefit from a change and to begin shaping an optimised feedback intervention/ dashboard. More details of the survey development and execution, including sampling methods, content, data collection logistics, data management and analysis will be provided in Chapter V.

In-depth interviews were held with key informants purposively selected from among both 'implementers' (the designers and implementers of the dashboards) and 'users' (the recipients of the feedback through data displays from dashboards). With these two sets of interviewees, I explored how information based on dashboard visualisations is used in practice, and the needs, barriers and contextual factors that influence their ability to take action based on the presented information. Two different interview tools were developed and pilot-tested prior to commencing the study.

The combination of quantitative and qualitative methods can be undertaken using integrative, sequential, and triangulation approaches [128]. I chose a sequential, linear data collection approach, largely for pragmatic reasons due to the COVID-19 pandemic restrictions. As presented in Table 3, I combined an exploratory and explanatory approach in three sequential steps. Please note, that although feedback implementers and feedback users were separated for the interviews and analysis, implementers can be seen as users too, for example by using data and feedback for their own purpose (e.g. tracking quality improvements for decision-making).

Table 3: Sequential data collection methods and steps.

Sequence	Data collection method	Rationale
Step 1	**Qualitative inquiry:** Primarily exploratory (and partly explanatory) in-depth interviewing with feedback/dashboard implementers.	To inform the survey on design considerations, usefulness and usability perceptions, expectations, information needs, and behaviour change, while also reflecting on implementers' perspectives to contextualise and re-evaluate design and implementation processes.
Step 2	**Quantitative inquiry:** Online survey	To investigate which information visualisations are effective and how dashboards understood through quantitative measures, also measuring the perceived ease of use, usefulness, and factors contributing to the utilisation of dashboards, and including neonatal KPI importance and presentation.
Step 3	**Qualitative inquiry:** Primarily explanatory (and partly exploratory) in-depth interviewing as a follow-up to survey with dashboard users.	To explain findings from survey and reflect on implementer interview, and to identify additional factors not captured by the survey to interpret the data, while also reflecting on broader user experience to understand which dashboard details work better.

The sequential data collection allowed learning from each step to inform the next. All steps were informed by the theoretical framework and literature as explained earlier. As seen from Figure 13, I then selected an integrative triangulation mixed methods design to draw on the strengths of both quantitative (numerical and generalisable) and qualitative (narrative and contextual) data, and to enhance the reliability and validity of the findings, following best practices [128]. Each data set was analysed and interpreted separately and brought together in the discussion in Chapter VII. This approach allowed the generation of nuanced findings and triangulation across data sets, meaning results from different methods were compared and combined to add depth and validity.

At the interpretation stage, the focus was to learn from both research methods results - interpretability testing and narratives about the overall user experience to provide a comprehensive understanding of how to optimise effective use of dashboards for QoC.

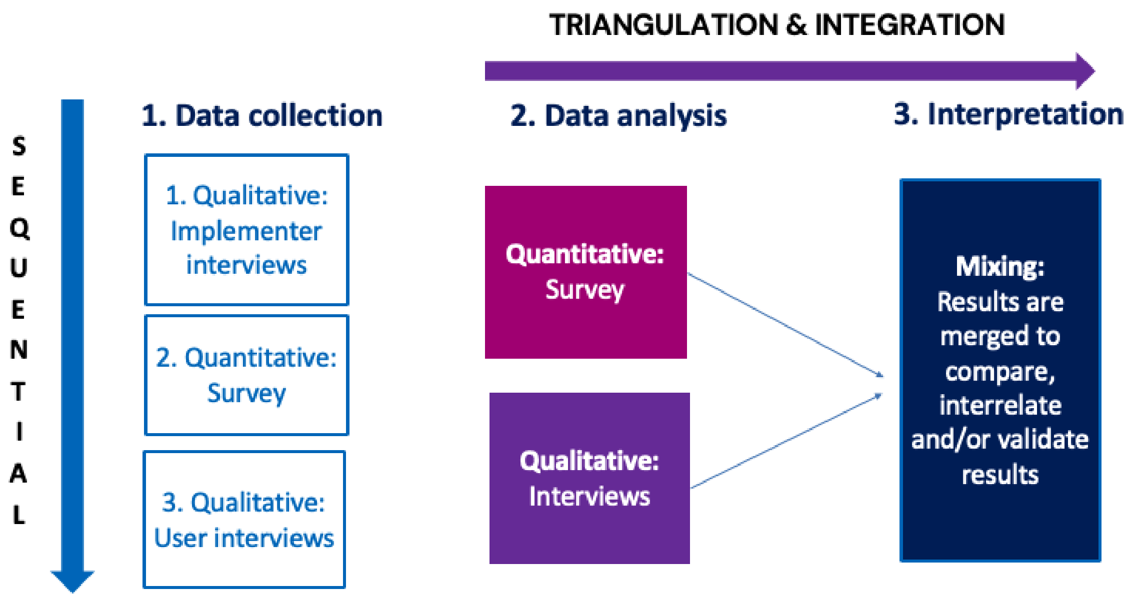


Figure 13: Overview of mixed methods data sets, sequential data collection and integration.

3.3.3. Research Team and Study Fieldwork

Fieldwork during this study had to be altered amidst the COVID-19 pandemic. I faced numerous barriers and delays, requiring adjustments to the research approach and data collection. A scoping review was added to complement the original design, and I re-organised my interviews to start with non-hospital staff (in my case the implementers). Due to unforeseeable travel restrictions, I had to reduce my time spent in Kenya from 12 months to 4 weeks. I had to begin my data collection entirely online and employ a local research assistant (RA) to assist with later interviews of hospital staff. To facilitate these changes, I had to submit a new ethics proposal. Luckily, I had prior work experience in Kenya and the health system and hospital context were familiar.

Finally, the research team besides myself, consisted of three supervisors and one experienced RA who at the time was a Master's graduate from Kenya employed by the KEMRI Wellcome Trust Research Programme (KEMRI) on a full salary. Additional senior KEMRI staff were involved as facilitators, alongside further facilitation from the Health Systems Collaborative research team and methodological advisory from a statistician employed by Green Templeton College in Oxford.

3.4. Setting and Collaborators

Here, I will describe the study setting in more detail, including the two programmes that implemented the two dashboards of focus in this study.

3.4.1. Kenyan Public Hospital Setting

Kenyan data and information systems were described in Chapter II. There are 6 service delivery levels (see Figure 14), where healthcare is provided, and data collected [129]. My research and data collection for this study is mainly concentrated on levels 4, 5, and 6 of Kenyan public hospitals, where a significant volume of maternal and newborn care is provided [130], and CIN and NEST dashboards implemented.

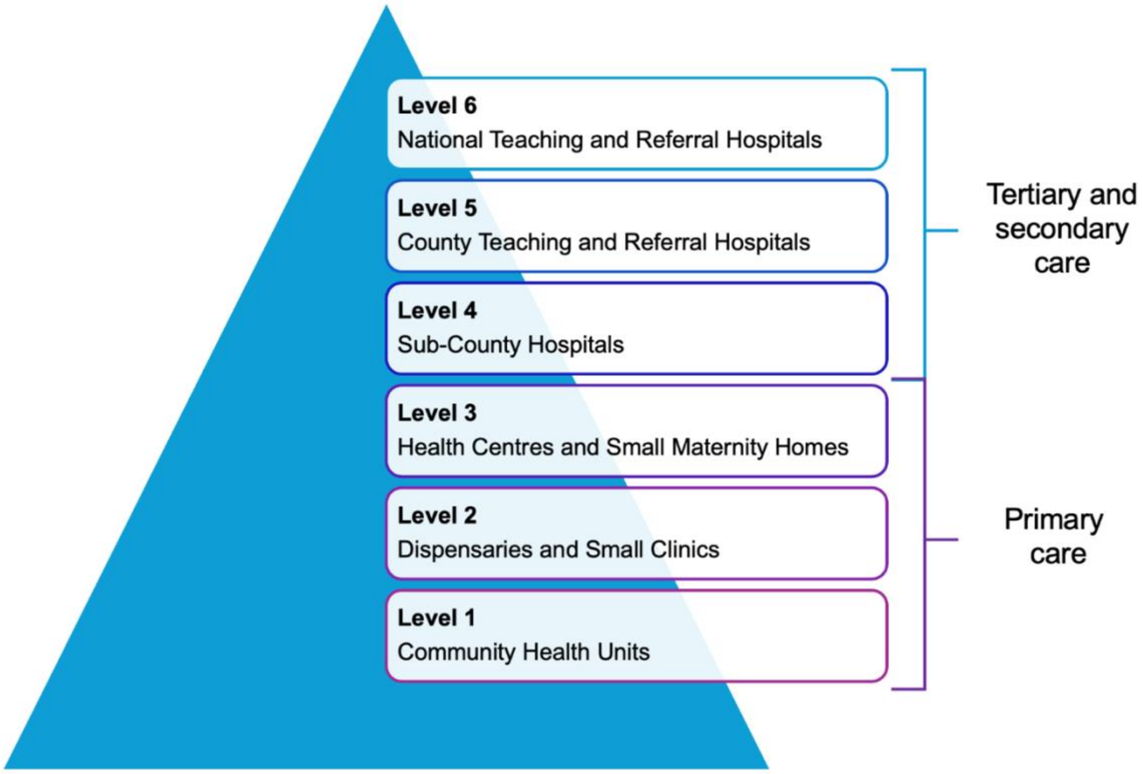


Figure 14: Levels of healthcare service delivery and hospitals in Kenya.

The dashboard implementers are typically external to the hospital environment, as they are responsible for the deployment, while hospital-based staff are the users. To answer research questions 2 and 3, fieldwork in the Kenyan public hospital setting was required. Identified target groups across the health system prior to starting the study were: Hospital staff (clinicians, managers, information officers), (sub-) county government officials, Ministry of Health officials~~,~~ and potentially also programme/project staff (those who are creating dashboards or employing them in feedback activities). Data collection began virtually (due to the COVID-19 pandemic and related travel restrictions) in October 2021 with dashboard implementers. However, data collection with hospital staff as research participants could only commence in April 2022 and was completed in September 2022. As noted above, more details about data collection are included in respective findings chapters.

3.4.2. Collaborators

I collaborated with two initiatives that have developed and deployed newborn performance dashboards: (a) the KEMRI's CIN (Clinical Information Network) implemented in Kenya (only) and (b) NEST (Newborn Essential Solutions and Technologies), implemented in Kenya, Malawi, Nigeria and Tanzania. Both initiatives work in partnerships with academic institutions within each country as well as universities from the UK and USA [22], [131]. Each newborn facility in public Kenyan hospitals, in which these initiatives have been introduced, employs an average of 20 clinicians (nurses and paediatricians) related to newborn and childcare, and about 10 health records information officers (HRIOs) and managers. And although both initiatives, CIN and NEST, introduced dashboards, they have never been formally tested to evaluate if they meet the actual needs of the users.

3.4.2.a. The Clinical Information Network (CIN)

KEMRI based in Nairobi and Kilifi, undertakes – in partnership with Kenyan public hospitals, the government and external research partners, such as the University of Oxford – significant clinical and health systems research aiming to improve newborn care in Kenya [132]. CIN brings hospitals with medical and nursing leaders from newborn units together to form a network in Kenya, as an intervention approach to improving service delivery across the country. Figure 15 shows where the CIN partner hospitals are located across different counties in Kenya and are listed in Appendix 3.1. During the time of my data collection, CIN was implemented at 20 public hospitals across 12 counties. However, since then, the network expanded its operations to 24 hospitals across 16 counties [131].

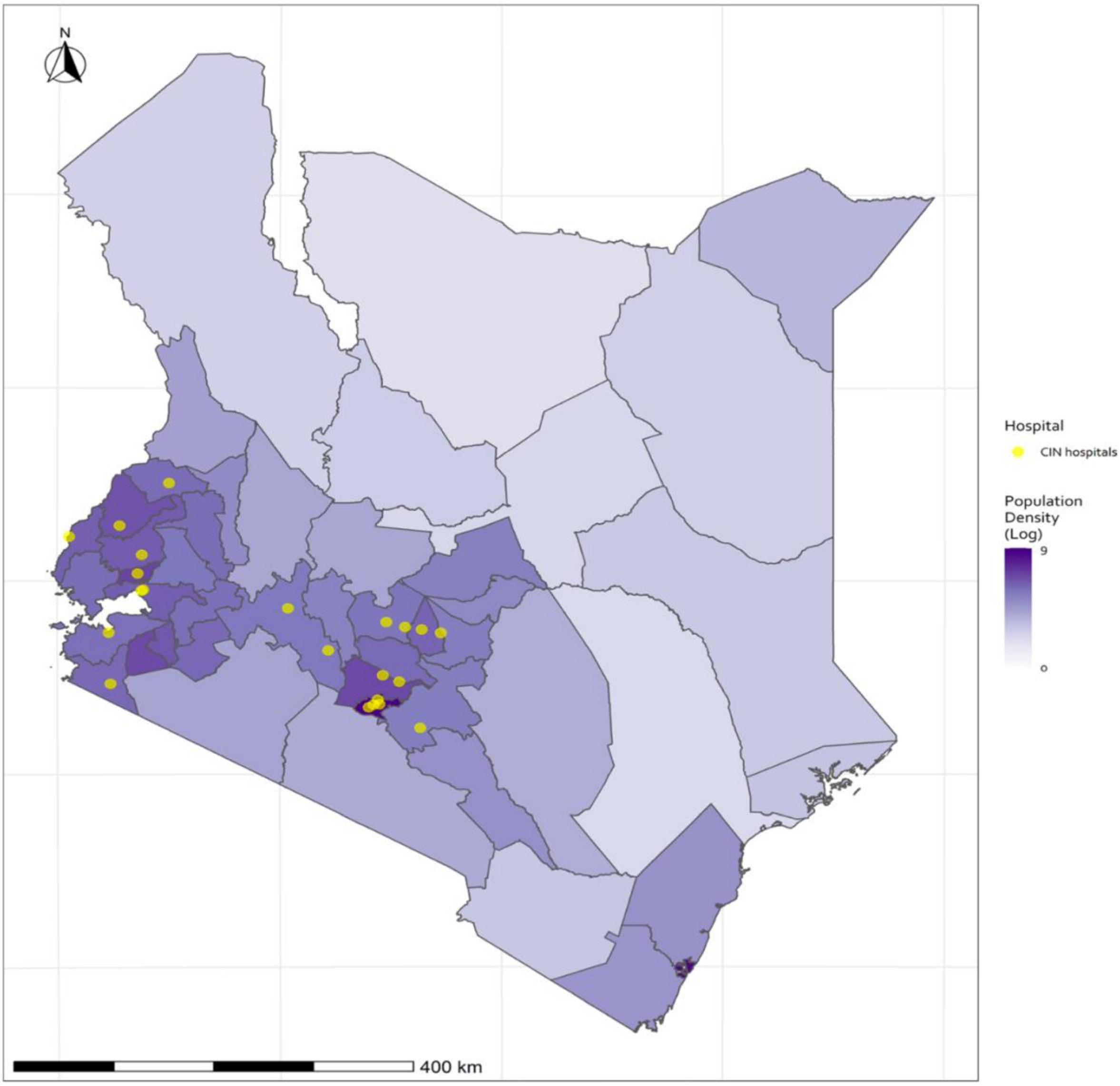


Figure 15: CIN footprint across Kenya, displayed on the country map. Darker shades indicate higher population density, yellow markers present CIN partner hospitals. Figure Source: BMJ Global Health, © English M et al. (2021) [133], used under cc by 4.0:https://creativecommons.org/licenses/by/4.0/.

As part of this initiative, CIN established a data collection system to enhance the quality and utilisation of patient data, and has expanded to include neonatal units, promoting better documentation and data-driven decision-making [31]. The idea was to use data to provide performance information to hospital staff on the ground and senior hospital managers, with a long-term aim to also use the same data to inform local government officials, who are responsible for, but not directly involved with the provision of clinical care. Currently, performance feedback on outcome and process indicators is primarily provided to partner hospitals on a monthly (shorter feedback) and quarterly (comprehensive feedback) frequency as seen in Figure 16 [117], [118].

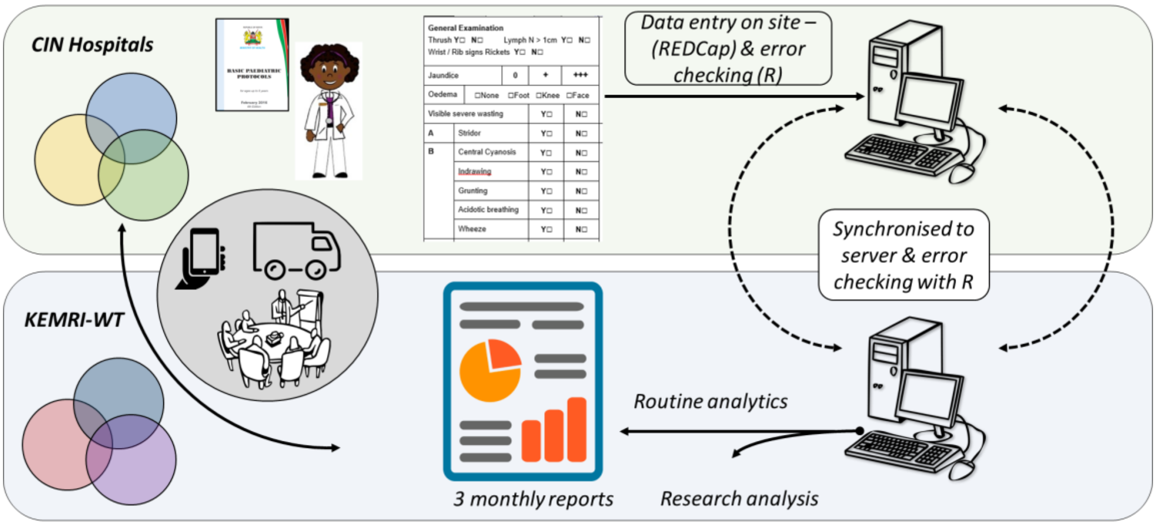


Figure 16: CIN Feedback system, © Ambrose Agweyu, KEMRI [131], used with permission.

At the beginning of this study, a fully digital dashboard had yet to be implemented. The feedback tool and process evolved from a paper-based to a hybrid tool, incorporating both electronic and paper-based feedback. Throughout this study I will use the term **dashboard** and refer to both the CIN and NEST tool.

3.4.2.b. The NEST360 Initiative

NEST is a multi-institutional partnership deployed in Kenya, Malawi, Nigeria and Tanzania committed to reducing neonatal mortality in these countries. As part of providing cost-effective, evidence-based technologies for neonatal care, NEST developed and introduced a dashboard for data-driven quality improvement at 13 partner facilities, all of which are also CIN partner hospitals [22], [116][[116]. At the start of this study in 2020, NEST had just designed and implemented the dashboard to track clinical care processes, outcomes of small and sick newborns, as well as the adoption and use of technologies they provide. Amongst other interventions (Figure 17) the dashboard contributes to quality improvement by providing real-time data feedback for data-driven decision making [21], [22].

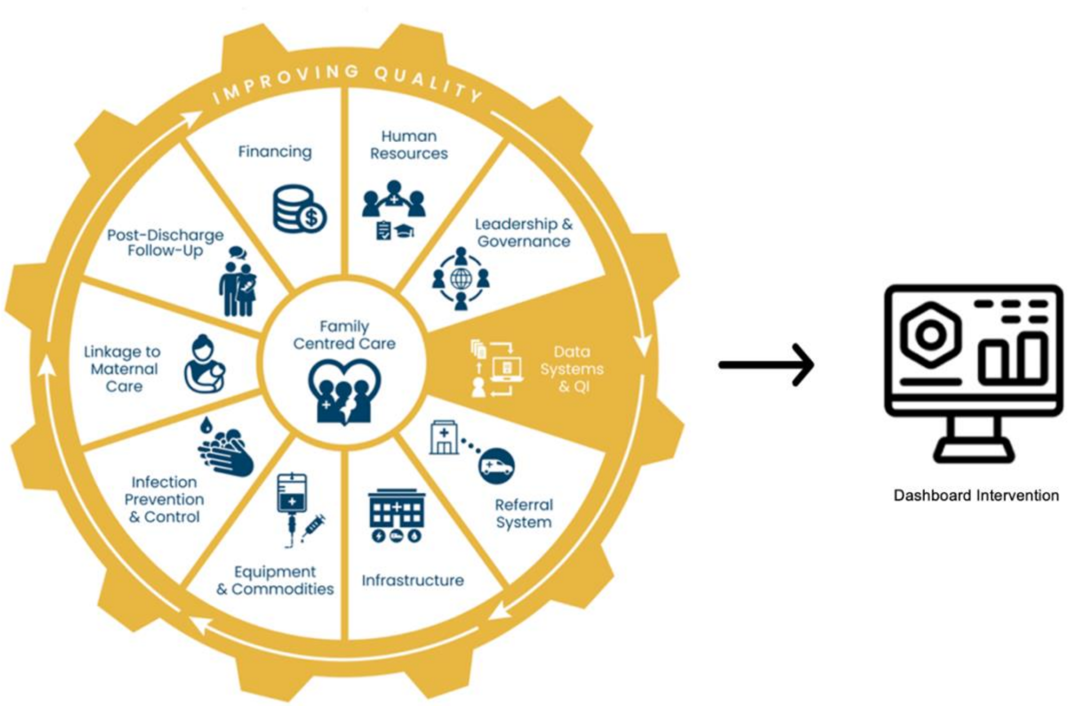


Figure 17: The NEST toolkit, presenting areas that the initiative addresses for improving QoC for newborns and of which the dashboard is part of. © NEST implementation toolkit [22], used with permission.

Thus, NEST added another tool to the existing CIN feedback system described above. More design details of the CIN and NEST dashboards are presented in Chapter V.

3.5. Positionality and Reflexivity

I acknowledge my positionality as a middle-class, German-native researcher with residency in the United Kingdom (UK) using English as a second language. As a white, cisgender woman with a higher education (Bachelor's and Master's level), my perspective is shaped by the social and institutional systems of both Germany and the UK. I have an academic background in Law, Public Health, and through this DPhil in Clinical Medicine and Global Health. My professional experiences include various roles in health informatics and public health consultancy, a clinical nursing internship in a hospital setting, and work with the World Health Organization. For the past six years, I have been involved in health systems & policy research, which involved field visits to hospitals in Kenya and Ghana. These diverse experiences provide me with an understanding of health systems and health service operations in daily reality although I do not possess the clinical skills of a physician. Not being from Kenya or another East African country means my research could not benefit from the language skills to communicate in Swahili and contribute with personal, culturally nuanced insights. This positions me as an outsider to the population I study. As neither a physician nor Kenyan, I rely on the research methods to capture people's voices, and I provide an external insight into systemic issues in newborn hospital care in Kenya with my research.

To mitigate this, I frequently engaged in conversations with my colleagues from Kenya (doctors and researchers both locally and abroad), which added further understanding of the local context, and to adopt a critical global health lens. I also engaged with both collaborators, in particular during the data collection and fieldwork phases. However, despite this collaboration, I sought to maintain a neutral stance to maximise objectivity and credibility when assessing both dashboards, avoiding any potential bias, and generating valid, evidence-based conclusions. This stance helped build trust for the user interviews, as I was able to position myself as an external researcher with no conflicts of interest, unaffiliated with either of the two initiatives and without direct employment with the collaborators or links to the Kenyan government.

Although I missed the opportunity to spend a longer period of time in the Kenyan hospital setting during this study, I spent some time in the setting prior to my DPhil research. I worked as a research assistant for one year on newborn data flows and health information systems in Kenyan hospitals, and I had the opportunity to visit a few CIN partner hospitals and talk with hospital staff between 2018 and 2019. Since I am a researcher without a clinical, but a legal and public health background, from a high-income country (HIC) in Europe, this work helped me understand the infrastructure, government policies and especially the clinical setting in Kenya, including the specifics of newborn units (NBUs).

When conducting health systems research on newborn quality care in an LMIC setting, I approach my work through a social justice, human rights, and intersectional lens, valuing equity and inclusivity. I question systemic power imbalances and advocate for locally driven solutions and ethical collaborations and aim to conduct research with sensitivity and to remain open to the diverse perspectives and experiences of my research participants.

3.6. Ethics and Research Integrity

When conducting my research, I adhered to the four ethical principles derived from Beauchamp's and Childless' (1991) principles of biomedical ethics [134]: respecting autonomy through informed consent, beneficence by maximising benefits, non-maleficence by minimising potential harm, and justice by treating all participants fairly and equitably. I followed all ethical guidelines, data storage and privacy guidelines set by the University of Oxford and the research partners, such as the KEMRI and NEST. Furthermore, I completed the research ethics training, as required by the University of Oxford (Appendix 3.2.), and adhered to best research practices aimed at ensuring methodological rigour and research integrity (reliability, credibility, replicability and reporting all possible limitations). To adhere to best practices, I used Equator toolkits and guidelines for quality and transparency of health research [135]: For the scoping review, I used the PRISMA Extension for Scoping Reviews (PRIMA-ScR) as a guiding tool [136]. For the qualitative component, I used the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist to ensure comprehensive and transparent reporting of the data collection and analysis [137], and the CASP (Critical Appraisal Skills Programme) checklist to critically appraise the quality and rigor of my qualitative research [138]. The use of these tools will be detailed in the relevant chapters and appendix.

Since my research was conducted with healthcare workers and partly in hospitals, ethical considerations must include sensitivity to staff shortages, ensuring that participation does not add undue burden to healthcare workers' time and responsibilities, especially during or shortly after the COVID-19 pandemic. This ensured that my research complied with all safety protocols and prioritised the well-being of hospital staff and patients during the crisis. As described in the COVID-19 pandemic impact statement, this research study experienced a nearly two-year delay, with data collection being paused and only resuming after all official national and international guidelines were lifted.

Furthermore, I conducted online interviews with implementers and reduced in-hospital walk-through interviews and observations to in-depth, talk-through interviews to minimise disruptions within the hospitals. Additionally, I reimbursed for participation following guidance by KEMRI, to acknowledge the time offered by hospital staff, without creating any influence or coercion.

For the empirical work, ethical approvals and permits for this study were issued by Oxford Tropical Research Ethics Committee (OxTREC 532-21) [139], and the Kenya Medical Research Institute (SERU 3459 and 4203). Since the study involved human participation outside of the United Kingdom, but did not include collection of human tissue samples or invasive procedures, it qualified as ‘minimal risk’ according to OxTREC and was approved on the 23 June 2021. All approvals are provided in Appendix 3.3.

CHAPTER IV: Scoping Review

This chapter presents the scoping review conducted to assess how dashboards have been used in neonatal care, and what can be learned about effective dashboard designs. It provides an overview of existing research publications on dashboard developments to track and present neonatal indicators in high-, middle- and low-income countries.

4.1. Introduction

4.1.1. Recap of Objectives

This scoping review synthesises the current evidence and knowledge that addresses an exploratory research question aimed at mapping types of evidence and is a useful method to identify how neonatal dashboards for QoC improvement can be developed and implemented. It answers the first research question (see Figure 11) by focussing primarily on how neonatal data is presented as information in dashboards globally, to learn best practices for developing effective newborn care dashboard designs. The review aimed to learn which factors, methods, design guidelines and KPIs worked when developing, implementing or evaluating dashboards using neonatal care indicators. It provides a broad understanding of how to guide future evidence-based dashboard development efforts by avoiding the repetition of less beneficial factors and help to inform the empirical part of this thesis. The expected outcomes of the review was to cover the three main themes of interest:

1) What are the development recommendations or guidelines for newborn care dashboards?

2) Which neonatal indicators have been selected to be presented with which graph types?

3) How were users engaged in the process to understand how they make use of the dashboard?

4.1.2. Review Relevance

Over the last decade, dashboards, scorecards and other visualisation tools have been increasingly released to track health indicators. The term ‘dashboard’ and ‘scorecard’ is often used interchangeably; however, scorecards are more structured and focus on organisational operations and have underlying management methodologies such as the Balanced Scorecard, Economic Value-Add (financial performance), and Six Sigma (process improvement) [3]. While dashboards “show how an organisation measured up against its goals, scorecards provide links to how goals support [a corporate] strategy”, and therefore, they can contain the same information but differ by the degree of knowledge and integration with processes [3], [140].

Information management tools that visually track, analyse or display KPIs, metrics and key data points are valuable and popular tools for translating data into readable information for feedback and presenting QoC performance indicators [25], [65], [143]. They are increasingly used for various purposes in healthcare and at different levels of health systems, such as tracking patient data, monitoring and evaluating health interventions, and providing feedback on performance to hospital staff and management [66], [67], [69], [71]. However, according to La Grouw (2012) “anyone can design a dashboard, but designing an effective dashboard is a combination of cognitive science and business process knowledge” [3].

While there is a lot of information available about dashboard design, implementation and use in healthcare and in particular in HIC, there is little information about how dashboards were designed and implemented using neonatal indicators specifically, and even less so in LMICs. Given the major focus on tracking and scaling up newborn care for SDG 3.2. [143], research assessing the tools to use this information is now very relevant. However, for dashboards to be effective, the dashboard aim, the specific setting and consequently user needs are important to understand to begin with.

4.2. Methods

Here, I present the scoping review design details, outlining the search strategy, inclusion and exclusion criteria, the selection of included articles, and data extraction.

4.2.1. Justification

With a scoping review I aimed to answer the first research question and provide a summary of existing research publications that have been conducted using dashboards to track neonatal indicators in any setting, including high-, middle- and low-income countries. The expectation was to learn which factors, methods, design guidelines and KPIs worked when developing, implementing or evaluating dashboards to improve quality of newborn care, and guide future evidence-based dashboard development efforts by avoiding the repetition of less beneficial factors.

I selected a scoping review to map a broad body of literature and include various study designs. Scoping reviews are a growingly common approach for synthesising research evidence by mapping broad topics [2]. Arksey & O'Malley proposed to define scoping studies as the method "*to map the key concepts underpinning a research area and the main sources and types of evidence available, especially where an area is complex or has not been reviewed comprehensively before*" [23]. This description has been extended by several colleagues, highlighting that scoping reviews provide the "*opportunity to identify gaps in research, and types of source of evidence to inform practice, policymaking, and research*" [24]–[26]. While scoping reviews and systematic reviews share many processes as they are both reviewing a body of literature in a rigorous, transparent and comprehensive way, the main difference is their purpose. As the purpose of scoping reviews is to map a body of literature, it is less concerned with finding the best available research, while systematic reviews often focus on research quality to synthesise empirical evidence from the least biased studies [3]. On the other hand, scoping reviews are more objective and comprehensive than narrative literature reviews. Scoping reviews do not evaluate the quality of studies but seek to provide an overview of a broad topic and include a greater range of study designs [25], [27], [29]. It is suitable to undertake a thematic construction and create a narrative to summarise results across the broad topic of identifying what dashboards are (and how they are) used for summarising newborn care performance. Finally, the results can be used as a first step to reinforce a systematic review when there will be more research publications with a specific focus on the design and development of neonatal care dashboards.

4.2.2. Search Strategy

A search strategy was developed prior to the literature search. With this strategy, I systematically searched published research articles from four databases: PubMed, Scopus, Web of Science Core Collection, and the Cochrane Library (Search Strategy, Appendix 4.1.). The search was limited to studies reported using the English language and a timeframe from (and including) year 2010 to 2020.

4.2.3. Study Inclusion Criteria

Eligible studies explored any information management tool that visually tracks, analyses or displays newborn KPIs, metrics and data points to monitor newborn health or specific newborn-related processes with electronic tools (Inclusion-Exclusion Criteria, Appendix 4.2.). Thus, eligible studies used dashboards, scorecards or similar electronic visual feedback systems. Definitions and terminologies of these tools were given in Section 2.3.1. All study designs and research articles from low-, middle- and high-income countries were included.

4.2.4. Study Selection

Maternity dashboards were considered eligible if they used at least some indicators specifically related to newborn care or alongside birth events. A total of 9,073 articles were screened, of which 3,512 were excluded as duplicates and another 5,487 as irrelevant or duplicates. I used Abstrackr, a web-based open platform for semi-automated citation screening, to assist in the title and abstract review phase [144]. Following the title and abstract review, two independent reviewers (CH, YZ) screened 74 full-text articles. The observed agreement rate was 93.24%, meaning both reviewers independently made the same decision in 93.24% of cases. Cohen's kappa was used to assess interrater agreement, accounting for chance agreement, using the following formula [145]:

$$k = \frac{p_o - p_e}{1 - p_e}$$

where Po is the observed agreement and Pe is the expected agreement by chance. The resulting kappa value was 0.81. Thus, 93.24% agreement and a kappa of 0.81 reflect strong interrater reliability, measuring different aspects of agreement. Disagreements (n=5) were independently reviewed and resolved by two supervisors (ME, CP). Any further disagreements were resolved by consensus among the review team (CH, YZ, CP, SM, ME). This process (Figure 18) resulted in 18 articles that were included in the review.

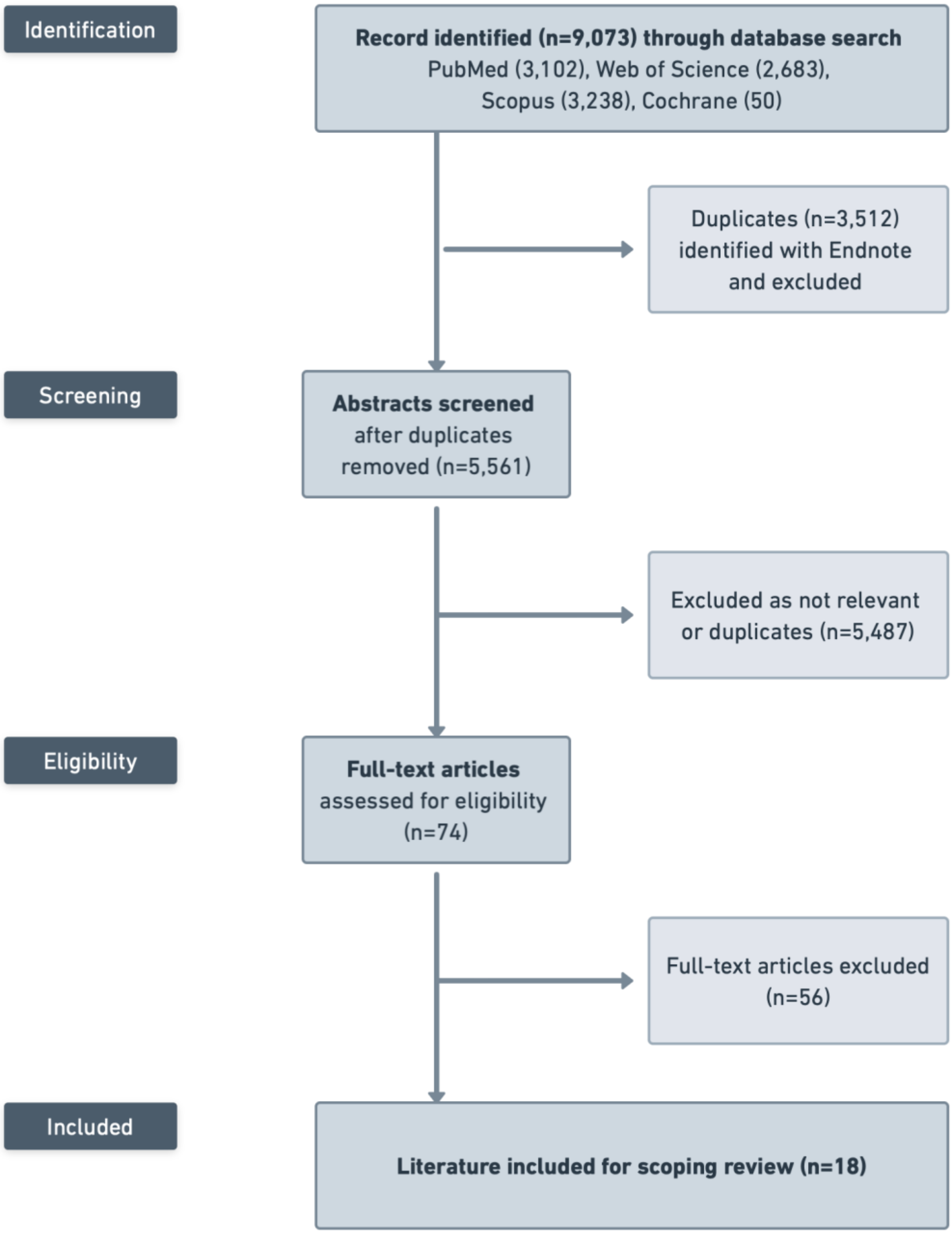


Figure 18: PRIMSA-ScR diagram presenting the screening, inclusion and exclusion of articles.

4.2.5. Data Extraction and Analysis

Results were charted in an iterative process for the three themes according to the objectives. Therefore, the information extracted for each study included: the study country, study design, service setting, study objectives, dashboard type, dashboard implementers and users, dashboard aim, data source or system, dashboard design and guidance followed, development and implementation processes, visualisation types and benchmarks, reported neonatal KPIs and how they have been selected, and finally whether any user experience analyses were conducted. Extracted data was reviewed by a second reviewer and disagreements resolved. Finally, I conducted a narrative synthesis of data to provide an overview of the identified literature and focused on the three themes.

4.3. Results

This section presents an overview of the identified evidence followed by a synthesis of findings regarding the three domains according to the research questions: the development and design of dashboards (theme 1), neonatal indicators and graph types (theme 2), and feedback of users (theme 3).

4.3.1. Overview of Global Evidence

Eligible studies (n=18) were identified (Overview of Studies, Appendix 4.3.) as mainly conducted by research teams from and for HIC (USA, Canada, UK, Republic of Ireland, Oman) and upper middle-income countries (Brazil, Colombia, Iran, South Africa). One study was conducted for a lower middle-income country setting (Bangladesh). No studies from low-income countries according to the World Bank classification were identified in this review [146]. The authors were looking at different study objectives but mostly related to developing, implementing or assesssing A&F and performance tools in the form of a dashboard or digital tool, and as part of that, to identify neonatal KPIs to measure maternal-newborn care. Some studies tackled more than one objective and investigated other important key factors before or after designing a dashboard and report these findings within one publication. For example Sprague et al. (2012) looked at which neonatal KPIs are suitable to measure quality, and how to provide automated A&F on these before investigating how to develop a dashboard. Dunn et al. (2016) examined not only the effect of implementing a dashboard but also explored differences in using it and measured factors of differential effectiveness, in addition to factors that predict differences in performance.

A variety of different study designs and methods were used: Mixed methods, surveys, systematic literature reviews, semi-structured interviews with dashboard users/ facility staff, a Delphi process, descriptive cross-sectional (case) studies and reports on the development and use of a dashboards, or an evaluation of a dashboard as part of other interventions. For most of the identified studies, the research teams (authors) were also the tool implementers (except for the District Health Information Software, version 2, hereafter: DHIS2) collaborating with hospitals, governments or insurance providers to improve hospital care.

The investigated setting was mainly public maternity hospitals, hospitals and academic centres with newborn units/services at different levels (facility, district, and national). Some studies were conducted as part of intervention programmes, such as Kangaroo Mother Care, Prevention of Mother to Child Transmissions (PMTCT), newborn screening, Antimicrobiale Resistance (AMR) surveillance, birth defect prevention and simulation training. Four published articles in Table 4 (No. 2, 8, 9, 13) relate to the same initiative from Canada, The Better Outcomes Registry & Network (BORN), which is a maternal, newborn and child registry to facilitate QoC for families across Ontario [147]. The findings from these studies, therefore, are complementary to each other. The digital visualisation tools used by the authors to present neonatal indicators were dashboards (n=14), a decision support tool (1), a visual analytics tool (1), a hospital report card (1) or an online platform (1)

Table 4: Overview of identified research articles

No.	Year	Country	Dashboard Category	Dashboard Type	Dashboard Aim	Authors	Title
1	2010	UK	Dashboard	Clinical performance dashboard	To analyse QI and provide feedback	Flanning C. et al. [148]	Dashboards in neonatology
2	2012	Canada	Dashboard	Maternal-Newborn Dashboard	To provide a clinical A&F system for QI	Sprague, A. et al. [149]	Measuring quality in maternal-newborn care: developing a clinical dashboard
3	2013	UK	Dashboard	Maternity dashboard	To monitor performance and quality of services	Simms, R. A. et al. [150]	Development of maternity dashboards across a UK health region; current practice, continuing problems
4	2013	Brazil	Dashboard	Maternal-Newborn Dashboard as part of the Obstetric EHR System	To provide a healthcare indicators dashboard	Gaspar, J. et al. [151]	Maternal and neonatal healthcare information system: Development of an obstetric EHR and healthcare indicators dashboard
5	2014	South Africa	Dashboard	Visual dashboard and data for action reports	To monitor PMTCT-related performance and QI	Bhardwaj, S. et al. [152]	Elimination of mother-to-child transmission of HIV in South Africa: rapid scale-up using quality improvement
6	2015	South Africa	Dashboard	Clinical performance dashboard	To monitor neonatal bloodstream infections and clinical performance	Raban, M. S. et al. [153]	Impact of an educational intervention and clinical performance dashboard on neonatal bloodstream infections
7	2015	The Nether-lands	Scorecard	Scorecard for newborn information as part of screening programme,	To detect neonatal seizures/ antenatal risks	Temko, A. et al. [154]	Clinical implementation of a neonatal seizure detection algorithm
8	2016	Canada	Dashboard	Maternal-Newborn Dashboard	To provide a clinical A&F system for QI	Dunn, S. et al. [155]	A mixed methods evaluation of the maternal-newborn dashboard in Ontario: Dashboard 8attributes, contextual factors, and facilitators and b9arriers to use: A study protocol
9	2017	Canada	Dashboard	Maternal-Newborn Dashboard	To provide a clinical A&F system for QI	Weiss, D. et al. [156]	Effect of a population-level performance dashboard intervention on maternal-newborn outcomes: an interrupted time series study

10	2017	Colombia	Visualisation tool	Visual analytics tool	To monitor, track newborn growth data	Diaz, D., et al. [157]	Visual tools for the exploration of growth data in a cohort of kangaroo infants during their first year of life and to identify the influence of variables
11	2017	USA	Dashboard	Maternity/ Birth defect Dashboard	To track and report dashboard shoulder dystocia	Szymanski, L. et al. [158]	Implementation of a multicenter shoulder dystocia injury prevention program
12	2017	USA	Scorecard	Hospital report cards for diagnoses information and visualisations	To monitor diagnoses performance and screening protocols	Patrick, C. et al. [159]	Improving analysis and reporting for Virginia's Division of Consolidated Laboratory Services newborn screening program
13	2018	Canada	Dashboard	Maternal-Newborn Dashboard	To provide a clinical A&F system for QI	Reszel, J. et al. [160]	Use of a maternal newborn A&F system in Ontario: a collective case study
14	2018	Iran	Dashboard	Dashboard as part of NICU Clinical Decision Support System (CDSS)	To design and evaluate an AMR surveillance system	Rezaei-Hachesu, P. et al. [161]	The design and evaluation of an antimicrobial resistance surveillance system for neonatal intensive care units in Iran
15	2018	UK	Dashboard	Maternity dashboard	To review and produce a list of indicators	Bunch, K. J. et al. [162]	Developing a set of consensus indicators to support maternity service quality improvement: using Core Outcome Set methodology including a Delphi process
16	2019	USA	Online platform	Online data visualisation and exploration tool	To monitor and relate preterm births to environmental factors	Jankowska, M., et al. [163]	An Online Geographic Data Visualization Tool to Relate Preterm Births to Environmental Factors
17	2019	Oman	Dashboard	Maternity dashboard	To improve QoC at ground level	Patel, M. S. et al. [164]	Development and Implementation of Maternity Dashboard in Regional Hospital for Quality Improvement at Ground Level: A Pilot Study
18	2020	Bangla-desh	Dashboard	DHIS2 dashboard not main purpose of study	To monitor services - dashboard not main purpose of study but a result of evaluating DHIS2	Begum, T. et al. [165]	Perceptions and experiences with district health information system software to collect and utilize health data in Bangladesh: a qualitative exploratory study

4.3.2. Newborn Dashboard Development

Within theme 1, I looked at how (newborn) dashboards have been designed and implemented, and maybe even evaluated or informed by specific guidelines. As far as studies described the development steps of their tool(s), they reported different approaches and recommendations, which are presented here.

4.3.2.a. Target Group and Data Sources

As a result of our inclusion criteria and search strategy, the main target group for the implemented dashboards was hospital staff, practices or midwifery services. Accordingly, target groups and users were mainly obstetricians, neonatologists, paediatricians, hospital or programme managers, midwives, nurses, technicians, system analysts and statisticians, researchers (only [163]) and DHIS2 users (only [165]).

Data for feeding information in analysed and visualised form to dashboards was mainly collected and provided from hospitals or insurance providers [158], and populated by a variety of data sources/systems, e.g. electronic maternity databases, electronic, paper-based patient records, registries, birth cohort database, DHIS2, with sometimes data being collated still manually. The BORN Some studies specifically set up information systems, and hospitals were asked to use them to feed the BORN Information System (BIS) with data [149], [155], [156], [160].

4.3.2.b. Reported Guidance and Frameworks

The identified articles were then screened to see if they used any previously published frameworks or guidance for developing, designing and implementing a dashboard in healthcare.

BORN studies investigated various data feedback mechanisms and tools (scorecards, dashboards, and standard paper-based reports). In a second step they develop a strategy for the dashboard implementation (following the process recommended by The Royal College of Obstetricians and Gynaecologists (RCOG) [166], which suggests that each hospital/unit requires clearly defined mechanisms and named individuals who are responsible for addressing issues as they arise). Plans were established for the post-implementation phase to test the effectiveness of knowledge translation strategies and to determine whether certain combinations of interventions will improve the uptake and use of the dashboard, and ultimately improve outcomes on the performance indicators [167]. Likewise, Simms et al. and Patel et al. refer to the guidance of the RCOG guidance and specifically to a guide for good indicators by the NHS Institute for Innovation and Improvement [150], [164]. However, both studies reported that the guide was mainly used for adapting parameters/ evaluating quality indicators and the recommended 'red-amber-green' (RAG) traffic light, together with using a performance threshold for displaying data.

Rezaei-Hachesu et al. referred to previously conducted descriptive studies by the same research team using a knowledge, attitude and practice (KAP) study to propose, design and describe the architecture of an AMR surveillance system, which has been the basis for a recommendation of an *"innovative architecture for the implementation of bio-surveillance systems"* [161]. The authors used this architecture for the pilot implementation of the AMR surveillance system in NICUs and used the Scrum methodology by Hicks and Foster for designing the system [168], [169].

Diaz et al. used Tamara Munzner's Visual Analytics framework for their visual analytics tool, which focuses on three factors: The What (data), the Why (tasks) and the How (visualisation idioms) [157], [170]. As the goal of this tool is to track newborn and child growth of premature babies, the authors also proposed a process model for problem domains to evaluate the effects of a specific event and the tool was designed to follow this model with a set of coordinated visualisations for each step (environment, pregnancy, birth, growth).

Jankowska et al. present with their study how to create an online data visualisation and exploration tool that can be used by researchers, clinicians and community health organisations to explore causes of preterm birth and potential interventions, highlighting the importance of user-friendly designs [163]. The design team used a user-centred design model and the user-utility-usability loop to design interactive maps by Roth, Ross and MacEachren, which prompted them to collect input and feedback on needs and designs from stakeholders and researchers (users) to revise the tool conceptualisation and functioning (utility) and finally to evaluate the new data visualisation tool version as a restart of the loop [171], [172], [173].

4.3.2.c. Dashboard Development Steps

Other studies reported the steps they have initiated to develop and employ their dashboard/tool. Gaspar et al. reported 7 steps for developing an obstetric EHR and healthcare indicator dashboard: 1) Variable need to be to compose a condensed obstetric EHR, 2) Values from manuscript to computer have to be transcribed, 3) Adequate maternal and neonatal indicators have to be identified and defined, 4) the specific software for inputs/data collection, storage and analysis has to be implemented and be capable of composing the obstetric EHR , and generate indicator panels, 5) the prototype installation needs to be tested in the facility and integrated to other existing systems, 6) results and impact of the quality indicators should be evaluated, and 7) report should be created and released so that they can help to improve the QoC [151].

The case study by Bhardwaj et al. described the dashboard development in three steps that have to be conducted: Step 1) Data review, meaning analysis of data completed for the key indicators in the PMTCT cascade; 2) Bottleneck analysis and prioritisation of actions identified during discussions reviewing the data and priority actions; and 3) Progress tracking by monitoring trends in the PMTCT cascade and dashboards [152]. However, they notably have not (yet) described how to co-design actively with users from the start.

As part of a shoulder dystocia injury prevention and simulation programme, Szymanski et al. created a dashboard to track performance [158]. Process mapping, gap analysis, and barrier analysis were used to identify the common barriers to implement the shoulder dystocia simulation programme successfully. A needs assessment was undertaken, and the simulator identified by the group consensus was provided to subscriber hospitals. A video and webinar presentation were developed to assist the hospitals with implementing the simulation programme. The metrics architecture was initiated with process measurement and clinical outcomes, as well as the metrics framework, and then developed with feedback from a consultant and finally confirmed by their Advisory Committee. The authors highlight that the architecture will be continuously improved.

4.3.3. Neonatal Indicator Presentation

For theme 2, the aim was to identify which neonatal indicators and graph types have been selected for and presented with the dashboard.

4.3.3.a. Indicator Selection Process

Overall, the studies showed a wide variation of which neonatal indicators were used, which is partly due to the different study objectives. It has to be noted that most of the studies are pre-dated to the 2020 publication of the WHO's standards for improving the QoC for small and sick newborns in health facilities, which outlined newborn indicator standards (input, process, outcome) [59]. However, some authors do not describe how they selected indicators [151], [153], while most other colleagues describe very detailed how or why the selected specific indicators and refer to some identified guidance or recommendations on how to identify the right indicators in general (KPI Selection, Appendix 4.4.). Flanning et al. state that due to the lack of standards (at the time) of what a dashboard should present for quality improvement, they recommend it should be based on what the individual unit needs to monitor factors considered as important, which for their dashboard in a UK newborn unit included to add occupancy, environmental cleanliness and mandatory staff training. Adding staff numbers helped to make a case for successfully employing 10 additional neonatal nurses, which improved care [148].

However, one of the BORN studies investigated how to approach an indicator selection before deciding how to provide feedback on indicators and consequently followed a 'rigorous development process' for selecting core quality measures to guide selection of the KPIs for their dashboard. Evidence summaries for each of the potential indicators were conducted. To establish benchmarks for the KPIs, they used a combination of peer-reviewed literature, existing clinical practices within Ontario, and advice from clinical experts. One of the BORN studies specifically focussed on identifying KPIs to measure quality within the maternal-newborn system and followed Janakiraman and Eker's criteria for good indicators in obstetrics [149], [174]. To meet the criteria that indicators are easy to define and observe, and based on existing or easily collected data, the BIS was used, and data elements reviewed by the BORN dashboard subcommittee. This committee undertook a survey based on Hermann and Palmer's framework [175] for selecting core quality measures. Patel et al. refer to the guidance by the NHS Institute for Innovation and Improvement for what a good indicator is: A measure to describe as much about a system as possible in an easy understandable way [164], [176]. The authors acknowledge that it is essential for an effective dashboard *"to use a limited number of meaningful measurements, in accordance with local needs".*

4.3.3.b. Indicator Presentation and Delivery

Almost all dashboards used traffic-lights or colour-coded systems or tables (Column 3, Appendix 4.4.). While most of the developers provided set visualisations, only the DHIS2 dashboard allows users to create their own data visualisations due to the technical condition and functionality of DHIS2. No standards or best practices were reported regarding which specific graph types (e.g. bar, line or pie charts) work best and are easily understandable for all users and different target groups. Most authors assumed that dashboards and graphs are useful and easy to interpret, while using pie and donut charts [148]. Most dashboards used real-time feedback (Column 4, Appendix 4.4.), although some studies also reported in-person feedback sessions (weekly or monthly) as helpful additions to digital visualisation tools.

4.3.4. User Feedback and Improvement Recommendations

Finally, objective theme 3, aimed to extract information on what designs work best and what can be improved to identify unanticipated consequences and implementation challenges from users' perspectives. There is a general recognition that developing and implementing a dashboard is labour-intensive in terms of gaining clinical engagement and timely collation of the data, often from multiple sources. Developing robust systems to populate a dashboard is similarly recognised to be very labour-intensive, even when data are in digital form, and there would be significant advantages to greater automation [150].

For improving user performance, implementers mentioned visual benchmark setting and comparison within and outside the facility as particularly helpful [149], [154], [155], [159], [160]. At the same time, they reported that there is a perception that the tool itself means an increased workload for users instead of helping them in their daily routine. While implementers report the dashboard was described as a useful communication tool by users, concerns were raised about the uncertainty of how to interpret dashboard indicators and related visualisations and how to take action from it [154]. As a recommendation to improve the perception and use of dashboards from a user perspective, more training and less information/ less presented indicators were mentioned [154], [156], [165].

Begum et al. (2020) followed a grounded theory approach at every step of their analysis to understand user experience with and acceptance of implementing DHIS2 for collecting, analysing and using data [165]. The authors primarily looked at the overall use of DHIS2, the facilitators and barriers of implementing DHIS2 and therefore not specifically at how to develop a dashboard. DHIS2 provides an integrated dashboard for feedback loops, and in this context, the authors reported that the dashboard helped DHIS2 to be compatible for monitoring indicators. Only after the introduction of data visualisations with the dashboard did data entry increase because the directorates at the national level demanded using the platform for their own decision-making. Since the DHIS2 dashboard can be accessed and graphs be created by users, the authors reported that users made up to 16,000 graphs across 64 districts, meaning *"people were trying to use it"* [165]. However, it is unclear how accurately these visuals were conducted and interpreted by users. To strengthen the health information system and improve reproductive, maternal, newborn, child and adolescent health, participants of this study recommended an integrated online dashboard that reports on indicators and automatically updates performance statuses at (sub-)district levels.

4.4. Critical Assessment

Here, I provide a short critical assessment of the findings from the review. In Chapter VII, a discussion brings these findings together with those from the empirical component.

4.4.1 Dashboard Development Guidance

The scoping review has shown that some of the dashboard implementers investigated best practices and recommendations on how to develop a dashboard for improving care. However, authors referred to different guidelines and not many, thus there seems to be no standard and only limited guidance for what sequence and what design works best for neonatal data-specific dashboards and the newborn care setting. However, a large body of literature is available on how to design A&F interventions, (digital) tools and dashboards in the healthcare setting. At the same time, it remains unclear whether and how visualisations and graph designs were created based on best practices, although they have existed for more than a decade as seen in Section 2.3.3.c.

4.4.2. Dashboard User Experience

Published dashboard studies with neonatal care focus rarely report on how users (e.g., hospital staff) experienced the tools but tend to assume the dashboard content is easy to interpret. Additionally, they provide insufficient details on how users were engaged in the design process before or during the implementation (co-design with users) or in the evaluation process after implementation (feedback from users). Most recognised the importance of involving the users, considering context and different settings and the complexity of improving QoC with dashboards. However, behavioural insights from users and testing of the dashboard designs with users might shed light on the underlying complexities for dashboards to be effective for care improvements from a user perspective.

4.4.3. Contextual Challenges

The scoping review findings confirmed that there are, independent of HIC or LMIC settings, many complexities and challenges in the healthcare setting that make it challenging to design and implement dashboards, such as the underlying data and information systems. However, fitting information tools into existing systems requires interoperability and data standards that can conflict with integrating tools adapted to contextual hospital needs [177]. In addition, the difficulty of using dashboards for quality improvement when staff is short (particularly nurses and senior vs junior doctors) was highlighted in multiple sources [148], [150], [164].

4.4.4. Study Limitations

This scoping review focused on the evidence available, thus which research articles have been published to describe how neonatal visualisation tools work best. Therefore, grey literature non-peer-reviewed publications and other sources were not included. While grey literature, such as reports and internal documents, can provide valuable insights, it is very challenging to capture these comprehensively. Given the scope of this review, it was not feasible to include all grey literature.

However, I acknowledge that there might be more research projects that developed and implemented dashboards with the aim of improving neonatal care that have not been published and not identifiable. Throughout my DPhil research, I occasionally checked relevant sources, such as the NHS webpages, for guidance on designing and implementing neonatal dashboards, but no specific guidelines were available at the time. Similarly, I searched Google for any guidance on newborn dashboards specific to LMICs and found no relevant guidance.

It is important to note that since the review was conducted (2020-2021), there have been some new dashboard developments. For example, the Royal College of Paediatric and Child Health developed a new data dashboard for newborn care in the UK as part of their National Neonatal Audit Programme [178], and the WHO has since published a maternal and newborn dashboard to track progress on SDG 3.2. globally [11], which provides instructions on its use but not on the design process, although I could not identify any research publications linked to these projects. While the review could have been more comprehensive by including grey literature and non-peer reviewed resources, the search strategy and results still provide a valuable foundation for the empirical study and will inform future research.

4.4.5. Update: New Literature Since 2020

The scoping review is based on the timeframe 2010 – 2020 and was completed in 2021. The publication process was interrupted when a short time window to collect data for the empirical part of this thesis was possible shortly after COVID-19 pandemic restrictions were lifted. Since then, some additional literature was published; a quick database search identified three articles that meet the inclusion criteria, hence updating the review was considered.

According to Cochrane, the decision on whether and when to update systematic reviews depends on several factors: whether the review question remains relevant, if new evidence has emerged that changes the conclusions, and whether methodological advancements demand an update [179]. Although the guideline is intended for systematic reviews, there is no reason apparent for why it should be handled differently for other review types, at least in terms of searching for new findings.

Aligning with Cochrane's best practices, I checked newly emerging publications and identified three articles (Appendix 4.5.). However, no additional or conflicting evidence could be taken from these articles to answer research question 1. Consequently, and as I did not conduct the same rigorous review process for the broader body of literature from 2021-2023, these articles were not included in the scoping review for this thesis. However, to ensure the publication of this review in a peer-reviewed journal, a systematic screening of additional articles will be conducted to include new studies from 2021 – 2023 and a protocol for the scoping review was published (DOI: https://doi.org/10.17605/OSF.IO/EVQZP).

4.5. Chapter Summary

Although the scoping review results provide limited information in terms of how to design and implement a neonatal dashboard in the LMIC setting, it helped to identify specific themes that the empirical part of this thesis should include for this study to contribute to the existing literature.

There were several gaps in the evidence around detailed guidance on developing and implementing dashboards for newborn care, and specifics related to the LMIC hospital setting, and how users were engaged. Although understanding users' experiences, having user-friendly designs and involving stakeholders are reported to be essential components for creating genuinely beneficial tools, it was not reported how this ideally looks like and how dashboard designs were tested for their usability.

There could be more guidance on most effective forms and frequency of feedback using evidence that already exists from the data and information visualisation literature and how to measure and improve interpretability with specific graph designs [74], [93], [100], [104], [180], [181], [182], [183], [184], [185], [186], [187], from human-computer interaction [181], and human-centred design principle [25], [26], [76], [86], [171], [188]. Whether a living document approach providing real-time feedback or periodic reports (monthly or quarterly) works best is still unclear.

The user experience of newborn dashboard tools for quality and performance improvements remains underreported, despite its importance, particularly for the LMIC setting and achieving SDG target 3.2. Effective implementation of these tools requires attention to global policies about newborn KPI standards, and best practices about quality improvement through information and well-designed information systems, especially for complex systems like in healthcare.

CHAPTER V: MIXED METHODS – QUANTITATIVE ASSESSMENT: DASHBOARD USE

In this Chapter, I present the results of the quantitative component of the empirical study. After a quick recap of the objectives, I present the survey design and how the literature informed the survey questions, what graph types were tested, the recruitment process of participants, and how the data was collected. The survey results are presented spanning the participants' demographics, wider perceptions and general preferences of tools, and the perceptions of the CIN and NEST dashboards in terms of their ease of use and usefulness, followed by interpretation success rate and what factors influence interpreting the designs successfully. Finally, I summarise the findings before presenting the qualitative component in Chapter VI.

5.1. Introduction

5.1.1. Recap of Objectives

The study's research questions 2 and 3 are answered using a mixed methods design as previously described in the methodology chapter. The survey mainly answers research question 2 (see Section 3.3.), providing quantified results on how well key health sector actors in Kenya understand and interpret information from neonatal dashboards and other data tools, and what their perceptions are towards these tools. The quantitative survey results will then feed into the findings from the qualitative part by being able to interrelate findings when interpreting both data set results (Figure 19). I present findings from the quantitative and qualitative components separately and integrate both for interpretation and discussion in Chapter VII.

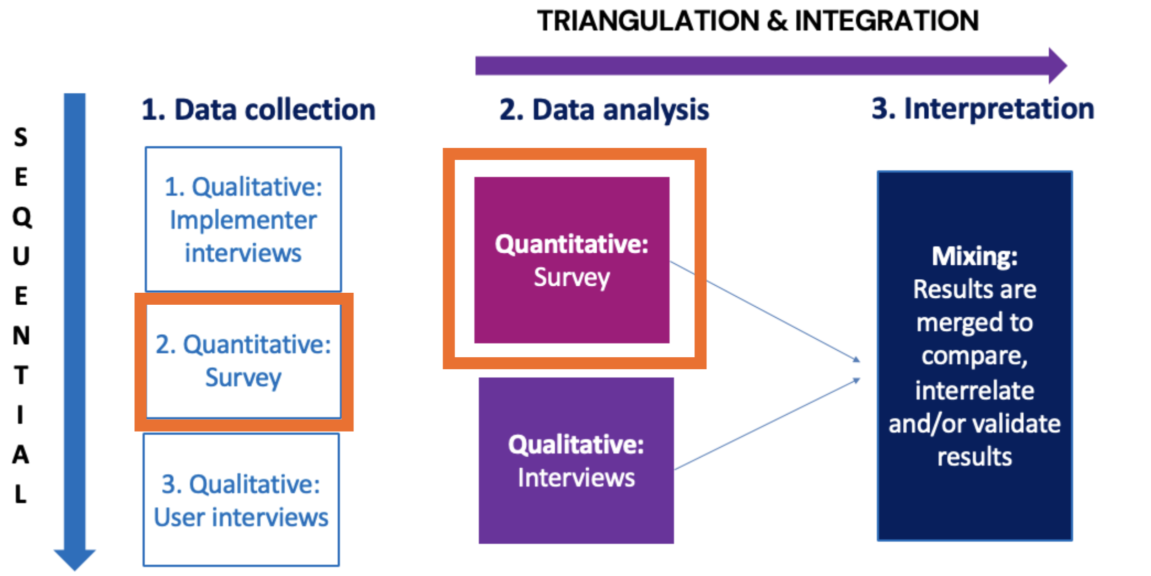


Figure 19: Overview of mixed methods data sets and the quantitative component presented in this Chapter.

5.1.2. Aims and Contribution of Quantitative Data

In the previous chapter, I reviewed the literature to learn from other evidence-based developments of hospital dashboards that are using neonatal care indicators and to answer research question 1. The 18 articles I reviewed showed few results in terms of providing step by step guidance of how to test or evaluate the success of a dashboards. Success in this context means that users can easily interpret and understand the presented information; hence, the tool is a valuable medium for translating data and driving action. Without published guidance on how to measure success, a mixed methods study was selected to test whether the tools for the CIN and NEST dashboard prototypes are useful for data-driven action and their usage by implementers and users. This chapter focuses on the quantitative survey as one component of the mixed methods study. The main aim of the survey is to learn what visualisation types work best and how dashboards are used and understood. Users' wider perceptions and other potentially influencing factors help to understand the context and eliminate confounders.

5.2. Methods

The survey was informed by the literature on how to collect quantitative data on the usability of the dashboards and evaluating feedback on performance. I led the selection of graphs, the recruitment of participants, variables of interest, data collection and the analysis plan with the support of KEMRI Wellcome Trust colleagues and my supervisors.

5.2.1. Literature Informing the Survey Design

The findings of the scoping review and the broader literature as outlined earlier informed the survey questions. In particular the work on assessing how 'useful' data and data tools were carefully developed with guidance from the A&F literature, the wider design, and data visualisation literature (Chapter II). The intention was to quantify users' perceptions of what they think is useful and easy to interpret as well as testing users on their abilities to interpret the specific graphs to correlate their perceptions with the actual task performance (interpretation success). In addition, I gathered data on other perceptions and contextual factors, e.g., other information systems in place, and confidence to improve QoC. All survey questions are provided in Appendix 5.1.

The quantitative part of my mixed methods study is important as it measures specific outcomes: in this context, the usability of a specific dashboard design in the form of quantifiable information from users on their experience of using the CIN and NEST dashboard prototypes and ability in interpreting information from these tools. I use the term 'usability' in my work with a main focus on the interpretability of dashboard contents and graphs. Usability testing broadly contains to evaluate how users can effectively interact with a system product or service, including how easily they can understand and use the presented content. Although I am not directly testing the interaction with the dashboards (as initially planned), I assessed key usability factors: interpretability (effectiveness), perceived ease of use (efficiency), and perceived usefulness (satisfaction)—which are central to usability as defined by standards (ISO 9241-11) [188]. This testing of a dashboard helps us learn what users would benefit from and what their needs and preferences are to inform newer designs with useful content while also helping us understand if dashboards can achieve their intended purpose, to change users' actions successfully. 'Successful' in this case means the user can make sense of the information that is provided and ideally use that information to improve performance. The success rates, as objective performance metrics, are important measures in addition to more subjective measures such as preference [189]. However, users' preferences and perceptions can help to understand the wider user experiences that contribute to dashboards being used [3], and hence, these were collected with Likert scale questions in addition. Collecting data on both, the survey results help to learn what the requirements of dashboards are for being useful tools to improve neonatal care.

5.2.2. Survey Design

The survey design is important to ensure high-quality, reliable data collection [191]. This section outlines the essential steps in creating a suitable survey to answer parts of research question 2, starting with defining which dashboard visuals were tested, the eligibility criteria and participant recruitment, the survey variables and measurements, and finally how the survey tool was reviewed and piloted.

5.2.2.a. Dashboard Graph Types

Since the CIN and NEST are used as dashboard prototypes for providing feedback on performance, specific design elements of how data have been visualised to readable information have been selected from those tools. Both tools analysed neonatal KPIs and created different graph types and forms of visualisations: Tables, simple line charts, multiple line charts, bar charts, stacked/grouped bar charts, and, in addition, used different colour codes, locations and frequencies of updating information. Different information displays from both tools were reviewed to extract graph types that are currently standardised in use to visualise data into information that provides feedback on neonatal data. The following graphs were selected and extracted from each tool for the user testing survey for CIN (a) and NEST (b):

i. Selection of CIN Graphs

CIN currently provides feedback through a hybrid feedback tool (as the precursor to becoming a fully digital dashboard) in the form of a multiple-page report that is circulated via email and presented during meetings using MS PowerPoint. This report is considered for my purposes to be a dashboard. Several CIN reports were reviewed and compared on their content and design to find the most updated standards that are used to provide feedback. Finally, I selected a report that was distributed at one of the larger referral hospitals in Kenya between July and September 2021. This kind of CIN report (Appendix 5.2.) was circulated in the same format to other CIN hospitals. The extracted graph types include: Survey *Graph 1*, a complex table with a colour-coded traffic light design to compare performance within a timeframe and with other facilities (Figure 20); survey *Graph 2*, a simple bar chart (Figure 21), and survey *Graph 3*, a line chart with performance measures shown in the figures below (Figure 22). A similar line chart was excluded from the analysis because the survey questions were unsuitable for calculating interpretation success (Appendix 5.3.).

		General Teaching and Referral Hospital A	General Teaching and Referral Hospital A	All Other Hospitals
		(Mar '21 – Jun '21)	(Jul '21 – Sept '21)	(Jul '21 – Sep '21)
Proportion documented	Total Cases	n= 493	n=375	n=6929
	Birth Weight	476/493 (96.55%)	362/375 (96.53%)	6786/6929 (97.94%)
	Admission Weight	464/493 (94.12%)	351/375 (93.6%)	4396/6929 (63.44%)
	Discharge Weight	368/493 (74.65%)	274/375 (73.07%)	5518/6929 (79.64%)

Color	Intepretation	Range
	Excellent Performance	>90%
	Good Performance	80 - 90%
	Some Performance	60 - 79%
	Poor Performance	<60%

Figure 20: Survey Graph 1: CIN Table presenting key performance issues of hospital A, comparing with previous performance and other hospitals' performance; used for survey questions 25-27.

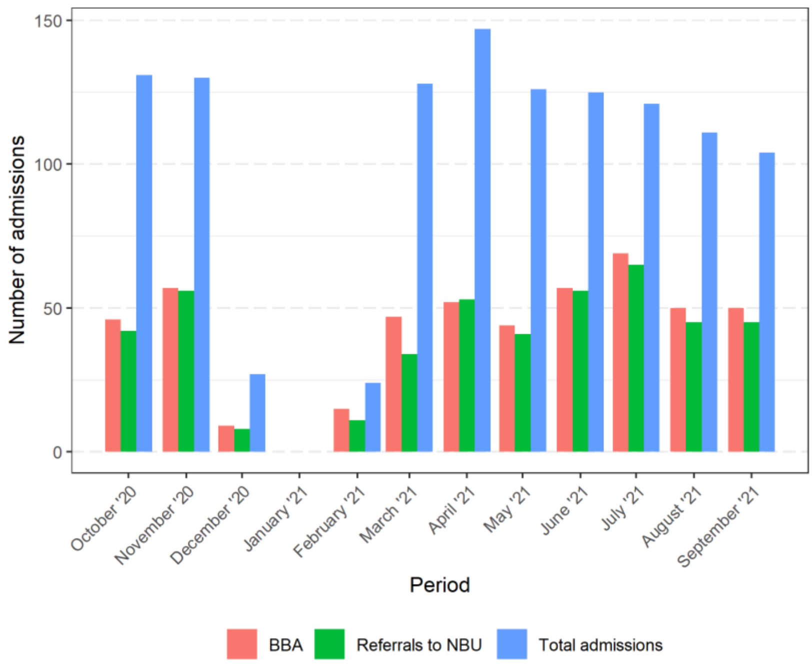


Figure 21: Survey Graph 2: CIN Bar Chart presenting admissions and referrals of newborns; used for survey questions 28 – 29.

Note: Referral into hospital newborn units (NBU) means those referred from other facilities. Born before arrival (BBA) means home deliveries or not born in any healthy facility. In how many months were referrals greater than BBA.

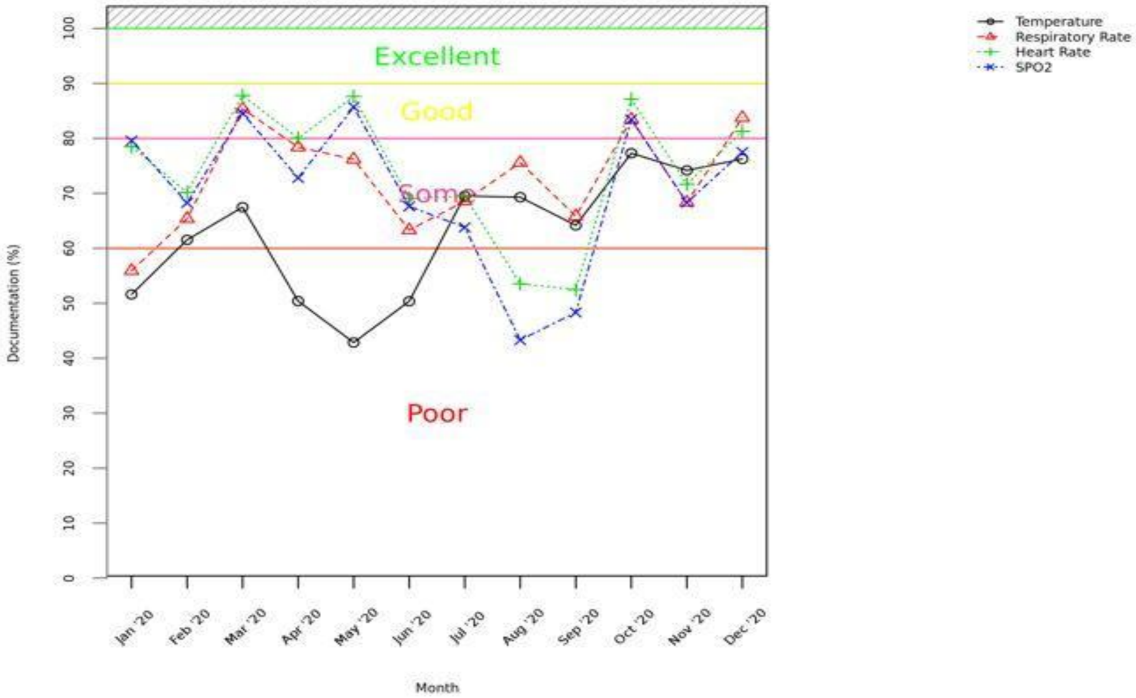


Figure 22: Survey Graph 3: CIN Line Chart presenting the documentation of signs and symptoms of severe illness; used for survey questions 32 – 33.

ii. Selection of NEST Graphs

At the time of data collection, NEST provided a one-page paper / digital feedback tool as a prototype of a planned fully digital dashboard. For the NEST dashboard, a version from January 2022, which shows all information from September 2020 until December 2021 was selected (Appendix 5.4.). The extracted graph types include: Survey *Graph 4*, the top part of the dashboard showing a line chart and a simple table (Figure 23); survey *Graph 5*, the middle part of the dashboard, presenting complex stacked bar charts, vertical and horizontal (Figure 24); and survey *Graph 6,* the bottom part of the dashboard, showing a simple table (Figure 25).

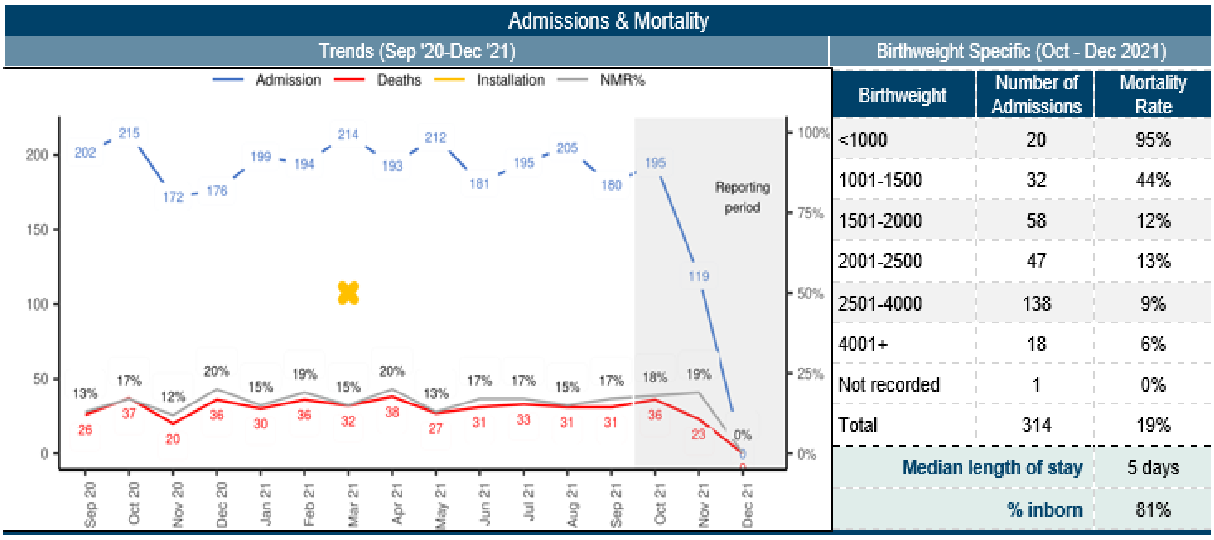


Figure 23: Survey Graph 4: Top section of the one-page NEST dashboard, presenting admissions and mortality trends; used for survey questions 34 – 37. Note: The questions and answers for the table on the right side were collated with the table responses for Figure 24.

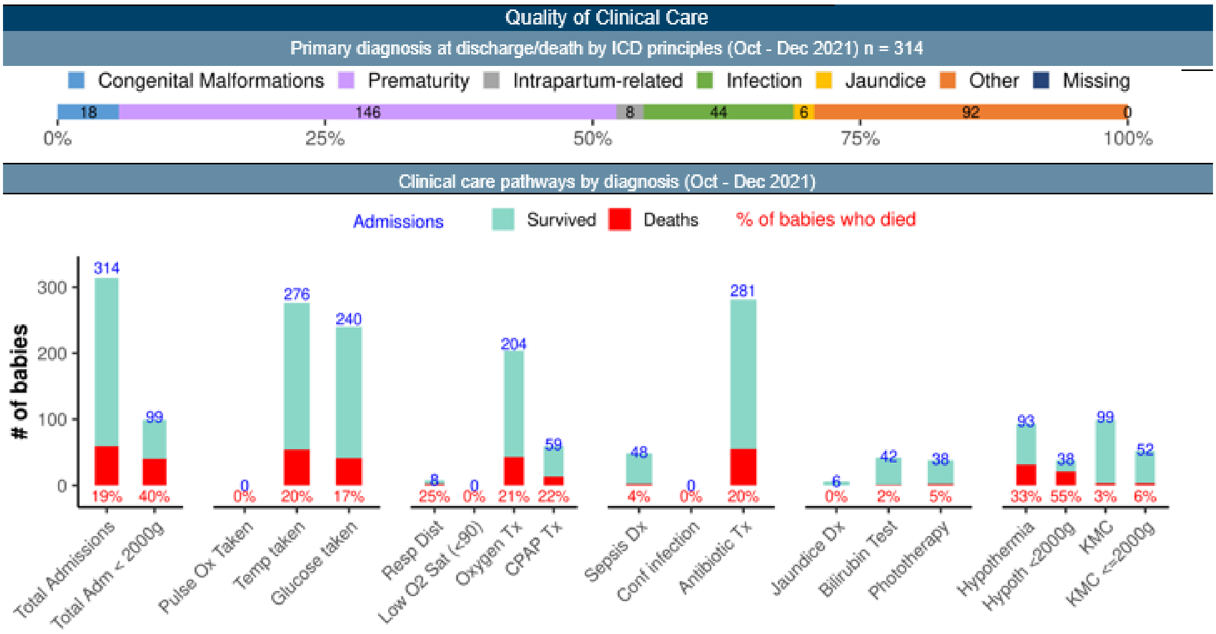


Figure 24: Survey Graph 5: Middle section of the one-page NEST dashboard presenting the quality of clinical care pathways; used for survey questions 38 – 43.

Health System Input as of Last QI or HFA Visit (Oct-12-2021)							
Human Resources			Device Functionality & Consumable Availability				
Personnel	Day of Visit	Night Before Visit	Oxygen		Warmers		Digital weighing scale
Clinical Officers	0	0	Concentrator	Tank	Radiant	Incubator	Functional: Not recorded
Nurses	5	2					Available: Not recorded
Doctors	2	1	Functional: 5	Functional: 4	Functional: 8	Functional: 12	Digital Thermometer
Biomeds assigned	Not recorded	Not recorded	Available: 5	Available: 4	Available: 8	Available: 12	
Day baby:nurse	48:5		Nasal prongs (neonate): Available		Probes: Available	Probes: Not available	Functional: 2
Night baby:nurse	48:2		Nasal prongs (infant): Available				Available: 2
Infrastructure (space, power, etc.)			Glucometer	CPAP	Suction Pump	Pulse Oximeter	Phototherapy
Bed capacity (normal)	65		Functional: 2	Functional: 5	Functional: 5	Functional: 3	Functional: 7
Bed occupancy rate	73.84615326		Available: 2	Available: 5	Available: 5	Available: 3	Available: 8
Power cuts >30min	Not recorded		Glucometer strips: Available	CPAP nasal prongs: Available	Catheters: Not available	Probes: Available	Lightmeter
							Functional: 1
							Available: 1

Figure 25: Survey Graph 6: Bottom section of the one-page NEST dashboard, presenting a table on health system inputs, human resources and devices; used for survey questions 44 – 46.

The selected graphs were chosen to investigate whether participants can interpret the content of dashboards and whether different designs and graphs (tables, line charts, bar charts, stacked bar charts) are easier to interpret than others for the specific setting. With those results, a ranking can be built on what graph type on which designs are easier or more difficult to interpret and explore whether this relates to what users say in their feedback about the tools.

5.2.2.b. Eligibility Criteria and Participant Recruitment

With the aim to test the tool for a specific setting, I identified the target population as users and potential users of hospital dashboards in Kenya. This included those users who might be partly working as tool implementers. Thus, I aimed to identify participants in Kenya from the 20 CIN and 13 NEST partner hospitals (with NEST hospitals a sub-set of CIN hospitals, as seen in Appendix 3.1.); the total population size of health professionals providing newborn care in these hospitals was estimated to be at maximum approximately 400 individuals (see Section 3.4.). Eligible participants could be any of the three cohorts: (Potential) users of NEST and CIN, NEST only users, CIN only users in Kenya. Furthermore, participants could be of any health system or hospital profession but focusing on nurses, paediatricians, health information records officers, data clerks and managers. To recruit participants, flyers with QR codes and online links were used to advertise the survey during CIN meetings in Nairobi and during the Kenyan Paediatrics Association Conference in Mombasa, in May 2022 (Figure 26).

Figure 26: Survey participant recruitment leaflet. Created by CH and advertised during the Kenya Paediatric Association (KPA) Annual Scientific Conference in Mombasa, 2022.

5.2.2.c. Survey Variables and Measurement

51 survey questions were developed to collect quantifiable information across 5 sub-research questions of interest. As seen in Table 5 below, the variables of interest included participant characteristics (1), wider user experiences (2), tool-specific user experiences (3), and how successfully the content can be interpreted (4), so that influencing factors on the success rate to interpret the content could be examined (5). The survey covered closed-ended and multiple-choice questions, as well as Likert scales.

Table 5: Overview of research question and sub-questions to be answered with the survey and collect information for the mixed methods study to assess the use of data with dashboards.

Research question 2: How well do key health sector actors understand and interpret visual feedback, and what are their needs (preferences) in order to act on it?		
Research sub-question	**Information of interest**	**Variables**
1. What are the participants' demographics and characteristics?	Participant characteristic	Gender, age, profession, user of digital devices, exposure/ user of feedback tool CIN/ NEST. (ordinal and nominal categorical)
2. What is the participants' wider experience and perception of information tools in use, feedback in general and QoC.	Wider perceptions and experiences	General preferences or perceptions about: other information tools in use, training needs, importance of neonatal KPIs, who are the main users, frequency of feedback, and confidence. (ordinal and nominal categorical)
3. What are the participants' perceptions of CIN and NEST feedback tools and their content?	CIN and NEST specific perceptions	Perceptions on how easy to use or useful (measured with Likert scales) the current design is. (ordinal categorical)
4. How successfully understandable/ interpretable is the content of CIN and NEST feedback tools?	Usability/ Ability to interpret the current CIN and NEST feedback tool design.	Interpretation success rates based on correct responses. (multiple-choice questions, scale)
5. Do participants' characteristics (1.), wider perceptions/experiences (2.) or graph design specific perceptions (3.) influence users' interaction with feedback tools (4.)?	Relationship: Do perceptions of how easy to use or useful it is, or other factors influence the ability to interpret.	Gender, age, profession, exposure, confidence, perception ease of use, perception usefulness (independent variable) and success scores. (dependent variable)

For research sub-question 4, aiming to test a user's performance, multiple choice questions were selected with correct and incorrect response options (1=correct, 0=false). For the 3 CIN graphs, respondents were asked 7 questions, while the 3 NEST sections had 11 questions. These questions were developed to assess respondents' ability to interpret a specific graph type and specific variations as they are currently used. By combining scores, I could evaluate the overall capability to interpret the CIN and NEST dashboard content, as well as all graphs combined. The total success score for both tools together, and for CIN and NEST individually, were calculated by averaging the success rate for each individual graph that has been tested with a minimum of one question, but in some cases with more than one question, e.g., if there was more content that could be tested from the graph. The results are, therefore, continuous as they have an infinite number of possible outcomes.

For the research sub-questions 1, 2 and 3, asking about participants' characteristics and perceptions and preferences as users, 33 closed-ended multiple choice and rating questions were used. Preferences of individual graph designs were tested after participants were asked to interpret the same graph type (user performance), assuming that they could only rate the design after they were exposed to having to use it.[190] To collect quantifiable results about user perceptions and preferences, Likert-Scale ratings from 1 (e.g., not useful) to 7 (e.g., very useful) were developed. The consistent use of scales from 1 to 7 was chosen considering that 7-points provide more nuanced results than 5-points but can also prevent survey fatigue in comparison to a 10-point scale.[192] The debates in terms of the analysis and interpretation of Likert items and scales (especially regarding the debate of the ordinal nature and non-normal distribution of Likert data, as well as the choice between parametric and non-parametric tests) have been carefully considered [193], [194], [195], [196]. However, in many cases using Likert scales parametrics and non-parametric tests perform equally well regardless of the data distribution [193], [194].

As recommended, results are presented as response frequencies (percentages for each category) for individual items and as a total score for the Likert scale (multiple items) [196]. Missing value handling was defined prior to the analysis. Missing values are treated as missing and coded NA. Statistics for each analysis are based on the cases with no missing or out-of-range data for any variable in the analysis as missing data was uncommon; thus, a complete case analysis was considered to be unbiased.

5.2.2.d. Survey Tool Review and Pilot

I set up the survey tool with 51 items using JISC Online (a platform recommended by Oxford University for data protection), and participants were able to access it via smartphones or laptops. All survey themes and individual questions were refined with a strong focus on the coherence of the themes of interest, including a careful consideration on: a) the quantitative vs. qualitative measures, b) the measurability with quantitative parameters, and c) the contribution they have to the qualitative parts of the mixed methods study (Section 3.3.). Two pilots were conducted prior to the data collection. The first survey pilot aimed to review typos, understandability of questions, order of questions and usability of the software. Three colleagues from the University of Oxford were invited and completed the exercise and subsequently asked for their feedback. They were not familiar with the CIN or NEST and could complete the survey within 30-45 minutes. The second pilot was conducted by inviting 9 Kenyan colleagues from the KEMRI Wellcome Trust, of whom 6 completed the exercise. The second pilot aimed to review understandability and the accuracy of context. Survey explanations and questions were finalised after these two pilots.

5.2.3. Data Collection Procedure

I collected the data during an organised CIN meeting in Nairobi and during the 3-day conference by the Kenyan Paediatrics Association (KPA) in Mombasa, in May 2022. For both, I arranged with the organisers to provide a 1-hour slot that I could use for introducing participants to the survey and give them the opportunity and time to complete to survey at that time, if they were happy to do so.

Participants were encouraged to use their own smartphones or laptops. Additional laptops and tablets were provided if necessary and a printed handout of the graph types for additional support was provided for participants to have the chance to look at graphs on paper if needed (Handout, Appendix 5.5.). Three researchers helped to facilitate the sessions and to assist respondents with practical aspects of completing the online survey. The initial time to complete the survey was generously scheduled to 50 minutes because the pilot test could be completed within 20-45 minutes.

Participants were informed about the study verbally at the beginning of the meeting, about their rights to withdraw their participation at any moment, as a well as with a written form at the beginning of the survey, where subsequently consent was collected (Participant Information Sheet, Appendix 5.6.).

5.2.4. Sample Size

In user experience research, there are varying recommendations and no consensus regarding the ideal sample size for usability testing [197], with some suggesting 3-5 users may be sufficient, while others suggest around 40 participants for quantitative studies [198] [199]. However, the lower sample size number recommendations come for practical reasons from product design iterations that consider cost-effectiveness and the necessity to fix companies' interface problems as quickly as possible. Since usability testing is considered the *"cornerstone of user-centred design practices",* researchers examined the impact of different sample sizes and could find surveys benefit from higher sample sizes. When testing how many problems can be detected, a sample of 50 users could find a minimum of 98% usability problems in 100 analysis samples [200]. For stand-alone surveys that compare designs based on binary metrics (such as success rates for graph interpretation) and rating scales (like Likert scores for preferences), both binary and continuous data types are often combined. This combination can lead to wider confidence intervals, indicating the need for a larger sample size to achieve low margins of error. Consequently, I calculated the sample size based on Cochran's 1963 formula for infinite populations for this study to ensure an appropriate 95% confidence level with an accepted 10% margin of error [201]. With Z = 1.96, p = 0.5, and E = 0.10 the sample size required for the study's aim was n = 96.04 as seen from the calculation here:

$$n_0 = \frac{Z^2 \cdot p \cdot (1-p)}{E^2} \qquad n_0 = \frac{1.96^2 \cdot 0.5 \cdot (1-0.5)}{0.10^2} \approx \frac{3.8416 \cdot 0.25}{0.01} \approx \frac{0.9604}{0.01} \approx 96.04$$

5.2.5. Data Analysis Plan and Software

The analysis plan (Appendix 5.7.). was outlined based on the objective to answer research question 2 and sub-research questions as presented in Table 5. Primarily, I aimed to establish proportions, and secondarily to explore factors expected to influence dashboard use and graph interpretability using measures of association. Statistical methods (descriptive analysis, frequencies, bi-variate correlations, multivariate analysis, ANOVA, t-test, Chi-square) were selected based on the type of variables that I collected. I extracted data as CSV files from the JISC survey tool and used Microsoft Excel and OneDrive for data storage. I used IBM SPSS Statistics (Version 29) for data analysis and R Statistical Software (Version 4.2.3; R Core Team 2023) for data visualisations [202], [203].

I preferred R Studio for data visualisations due to its flexibility and customisation options, enabling detailed graphics. Descriptive analyses of baseline characteristics and perceptions measured on a 7-point Likert scale were conducted with relative frequencies and correlations for all nominal and ordinal data (findings in Section 5.3.1. – 5.3.3.). Success scores for interpreting graph types were measured using correct or false response options. Factors contributing to these success scores were analysed using bivariate and multivariate analyses with (post-hoc) tests to explore associations statistically. These methods helped identify and understand the different variables influencing the success scores (findings in Section 5.3.4. – 5.3.5.).

5.3. Survey Findings

5.3.1. Participant's Baseline Characteristics

This section focuses on the results of sub-question 1 (see Table 5) to show participants' demographics and baseline characteristics, as well as exposure to the dashboards, taken from the responses to the survey tool questions 1 to 5 and questions 48 to 50 (Survey Outline, Appendix 5.8.). Using the approach of survey data collection meetings described previously, I could collect responses from 103 participants without any dropouts, representing approximately 25.75% of the estimated population of 400, indicating that a substantial proportion of the population was included in the study. 102 respondents worked in Kenya where the meetings were held, while one respondent was a visitor from the NEST programme in Malawi. The participant from Malawi had working experience in Kenya and was therefore included in the study. Respondents' characteristics are presented in Table 6 below.

Table 6: Demographics and baseline characteristics of participants.

	N (%)	Grouping/ Coding
Total number of respondents	103	
Gender		
Male	33 (32)	
Female	69 (67)	
Prefer not to say	1 (1)	
Age groups		
20 or below	0	Group 1: Younger age group, n=66
21-30	20 (19.4)	
31-40	46 (44.7)	
41-50	22 (21.4)	Group 2: Older age group, n=37
51-60	12 (11.7)	
61 -70	2 (1.9)	
71 and above	1 (1)	
Country of occupation		
Kenya	102 (99)	
Malawi	1 (1)	
Occupation		
Clinical Researcher	1 (1)	Group 1: Researcher/ Scientist & Others, n=12
Education Lead	1 (1)	
Researcher/Scientist	8 (7.77)	
Lecturer	1 (1)	
Programme Manager	1 (1)	
Hospital Manager	2 (1.94)	Group 2: Information Professionals, n=36
HRIO	31 (30.1)	
Data Clerk	1 (1)	
Data Manager	2 (1.94)	
Clinical Officer	1 (1)	Group 3: Clinical Professionals, n=55
Doctor	2 (1.94)	
Consultant	32 (31.07)	
Nurse	19 (18.45)	
Student	1 (1)	
Occupation in neonatal care		
Not related to neonatal care (Group 0)	5 (4.9)	Group 1: None to some NN care occupation, n=17
Partly related to neonatal care (Group 1)	12 (11.7)	
Directly related to neonatal care (Group 2)	86 (83.5)	Group 2: NN care occupation, n=86
Respondents are users of tool		
CIN feedback reports, yes/no	81 (78.64) /22 (21.36)	Group 1: Exposed to one of the tools, n=37
NEST dashboard, yes/no	49 (47.57) / 54 (52.43)	Group 2: Exposed to both tools, n= 47 Group 3: Exposed to none, n=19

Of all the 103 respondents, 47 respondents confirmed they were exposed to CIN and NEST dashboards as ‘users’ prior to the survey, 37 were exposed to only one of the two tools, and 19 reported they were never exposed to either of the two dashboards. The 81 respondents who confirmed they were users of the CIN dashboard all perceived the tool as useful (n=81). In the same way the 49 respondents who confirmed they were NEST dashboard users plus one who is not a tool user perceived the tool to be useful (n=50).

To understand whether respondents were in general familiar with and frequent users of technology and digital tools for their work, I asked whether they use a computer, smartphone and/or WhatsApp. All 103 participants use smartphones daily or at least a few times a week, 90 participant use computers, and 92 participants use WhatsApp in the same frequency. The variables on general familiarity of using digital tools, which essentially indicated universal and quite frequent exposure, was not used for any further analyses.

For further analyses, I merged reported occupations into three professional groups based on common characteristics and to yield potentially informative occupational subsets: 1. Researcher/ Scientist & Others (n=12), 2. Information Professionals (n=36), and 3. Clinical Professionals (n=55). In addition to asking participants whether their work is directly related to neonatal care, all participants were asked whether their occupation involves either patient care or patient quality improvement. Only 2 of 103 participants responded that their work does not involve any link to patient care or patient care quality improvement.

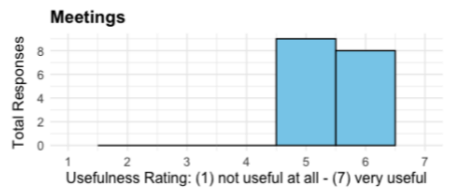
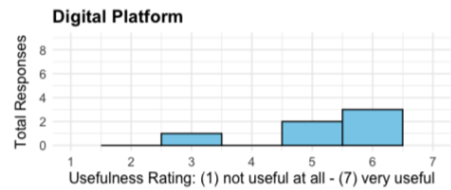
5.3.2. Wider User Perceptions

This section focuses on participants' wider experiences and perceptions of other health information tools and feedback in general to answer sub-question 2 (see Table 5). The data is taken from the responses to the survey tool questions 6 to 27 and question 51 (Survey Outline, Appendix 5.8.).

In Kenyan hospitals, all types of information systems and feedback types, such as registries, meetings, digital platforms, health management information systems (HMIS), and paper-based project reports, are used to retrieve information and/or to manage care. Participants therefore were asked to provide their perceptions of the usefulness and preference of these tools. Furthermore, the survey asked participants about perceived training needs, perception about which neonatal KPIs are most important for feedback, who they think the main dashboard users are, and what delivery frequency they prefer for receiving feedback. Finally, participants were asked to rate their own confidence linked to five QoC behaviour change variables. All perceptions were measured with a 7-point Likert scale with an ascending increase of preference, e.g., I don't like at all (1) to I like very much (7).

5.3.2.a. Interactions with Other Data Tools

In this section, I explore the tools that users primarily use in their work environment, their perceptions of the usefulness of these tools, and which tools they prefer over others. This provides insight into user preferences and the functionality of various tools in their professional context as presented in Table 7 below.

Tool type	a.PREFERENCE Total number (and %) of respondents (n=103) preferring this tool	b.USE Total number (and %) of respondents (n=103) mainly using this specific tool type	c.USEFULNESS Median *Mean, (SD)* of usefulness score amongst those using the tool	d. USEFULNESS Histogram Plots with rating Scale from 1 – 7 on X-axis, (1) not useful at all to (7) very useful, measured for total responses on Y-axis.
1. Patient Registries	9 (8.7%)	24 (23.3%)	6 *5.73, (2.31)*	
2. Health Management Information Systems for example DHIS2	15 (14.6%)	24 (23.3%)	6 *5.79, (1.58)*	
3. Meetings Virtual or in-person	21 (20.4%)	23 (22.3%)	6 *5.87, (0.8)*	
4. Digital platform that people can log on to, e.g. dashboard, website	56 (54.4%)	15 (14.6%)	6 *6.27, (1.12)*	

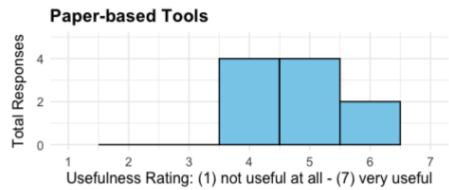
5. Paper-based reports including performance summary	0	15 (14.6%)	6 *5.53, (1.20)*	

Table 7: Tools in use, perceived usefulness and preference, measured with a rating on a Likert scale from 1 (not useful at all) to 7 (very useful).

Note: Not reported in this table is that there was 1 response for “My role does not involve performance indicators based on healthcare”, and 1 response for “None of these exist where I work”. In feedback field, one user mentioned that there is a “need for interpersonal interaction to complement digital platforms”.

56% of the respondents prefer using digital tools, while currently only 15% of the respondents use them as their main tool (Table 7, column a versus b). None of the respondents preferred using paper-based project reports with performance summaries if they had a choice to pick one of the named tools. Most used tools currently by this cohort were health management information systems, such as DHIS2 (n=24) and patient registries (n=24), while not necessarily preferring the latter (n=9). However, most participants perceived all tools as very useful tools with an average on the scale from 1-7 ranging from 5.53 to 6.27 and a median of 6 for all tools.

5.3.2.b. Training Needs

Here, I analyse the extent to which users feel they require additional training to use these tools effectively. Based on the tools in use mentioned above, the survey measured perceived training needs (a) to use the tool that is used most by a respondent, (b) to interpret data, graphs and tables, and (c) to use feedback to change practice and QoC (see Figure 27). The perceived usefulness was linked to the tool that participants indicated as the tool most used by themselves on a Likert Scale from 1 (I don't need more training) – 7 (I need much more training). The overall perceived training need (d) was calculated as the average (mean) of responses on (a) – (c) as shown in Figure 28.

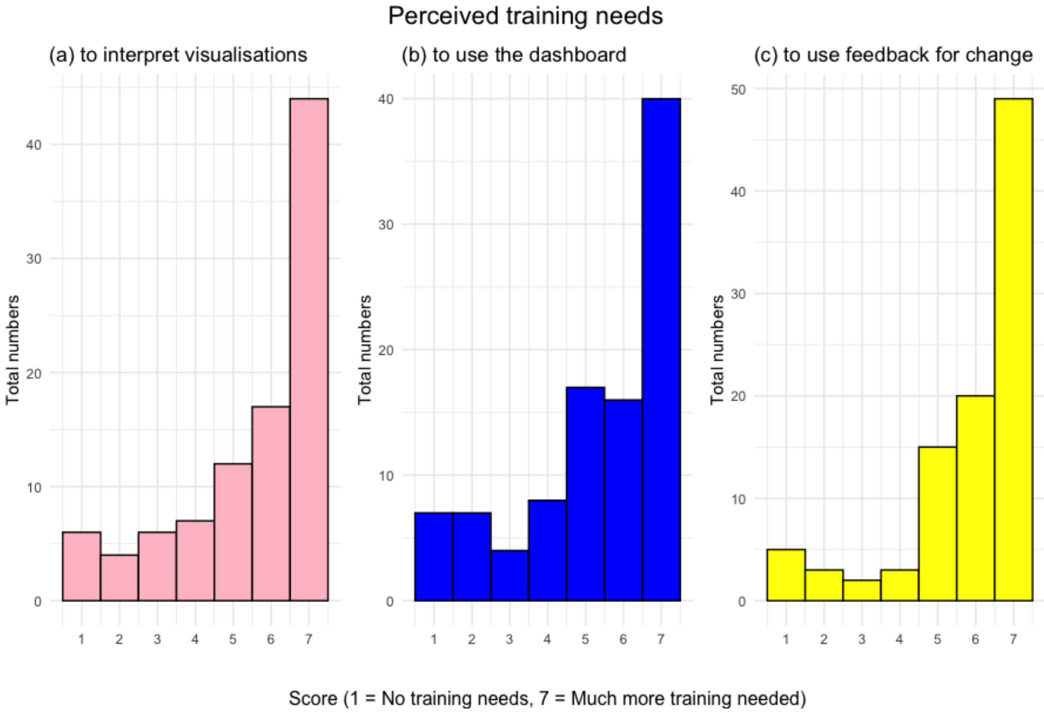


Figure 27: Training Needs Histograms – Perceived training needs, with 1-7 point Likert scale rating on x-axis with 1 (I have no training needs at all) to 7 (I need much more training) and total numbers of responses on y-axis, referring to training to (a) use the tool, (b) interpret content of the tool, (c) use feedback from tool for change in practice.

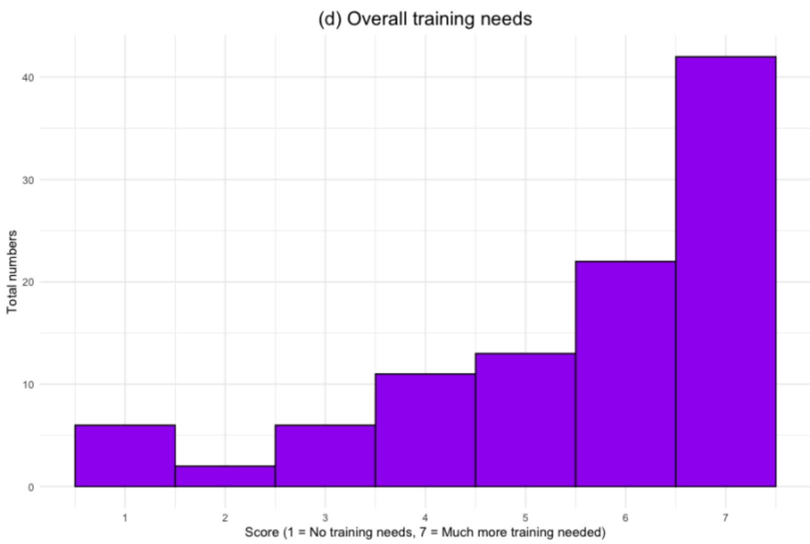


Figure 28: Training Needs Histogram – Overall training needs, with 1-7 point Likert scale rating on x-axis with 1 (I have no training needs at all) to 7 (much more training needed) and total numbers of responses on y-axis.

Note: This is the mean of the perceived training needs to (a) use the tool, (b) interpret content of the tool, (c) use feedback from tool for change in practice.

As seen from Figure 28, the majority of participants perceived that there is more training needed for all three categories (a-c). The consistency of Likert responses across the three response items was explored using factor analysis and the Cronbach's Alpha was .90, suggesting a strong between item correlation.[204] This suggests that a single (new) variable (d), which explains more than 83% of the variance, was suitable as a summary measure of training needs while the Kaiser Meyer Olkin test was 0.75, which determined that the sample size was sufficient for factor analysis.

5.3.2.c. Importance of Neonatal KPIs

Here, we focus on the participants' perceptions regarding neonatal KPIs. This section investigates which KPIs users consider most crucial to receive feedback on, providing a comprehensive understanding of the priority areas for feedback dashboards within neonatal care. From the selection of the following outcome and process indicators, all participants perceived them as 'important' for feedback on performance tools to improve the quality of neonatal hospital care:

- Number of Facility Births
- Stillbirth Rate
- Pre-discharge Mortality Rate
- Mortality Number of Newborn Deaths by Cause
- Number of Newborns Admitted for Inpatient Care
- Use of Vitamine K
- Use of Chlorhexidine
- Success of Early Kangaroo Mother Care for Premature Babies
- Success of Continuous Positive Airway Pressure (CPAP) Therapy
- Birthweight Always Documented
- Early Initiation of Breastfeeding Performed

Respondents could indicate their perception on a Likert scale from 1 (not important at all) to 7 (very important) or rate the indicator as irrelevant. The option 'irrelevant' was always provided as a separate 8th response option. Relative frequencies were calculated for the 11 indicators and correlations tested. Histograms for all indicators showed a highly skewed outcome towards the answer 'very important' for the KPIs included in the CIN and NEST dashboards, suggesting that the dashboards potentially deliver important information for practice (see Figure 29). None of the respondents considered any of the KPIs as 'irrelevant'.

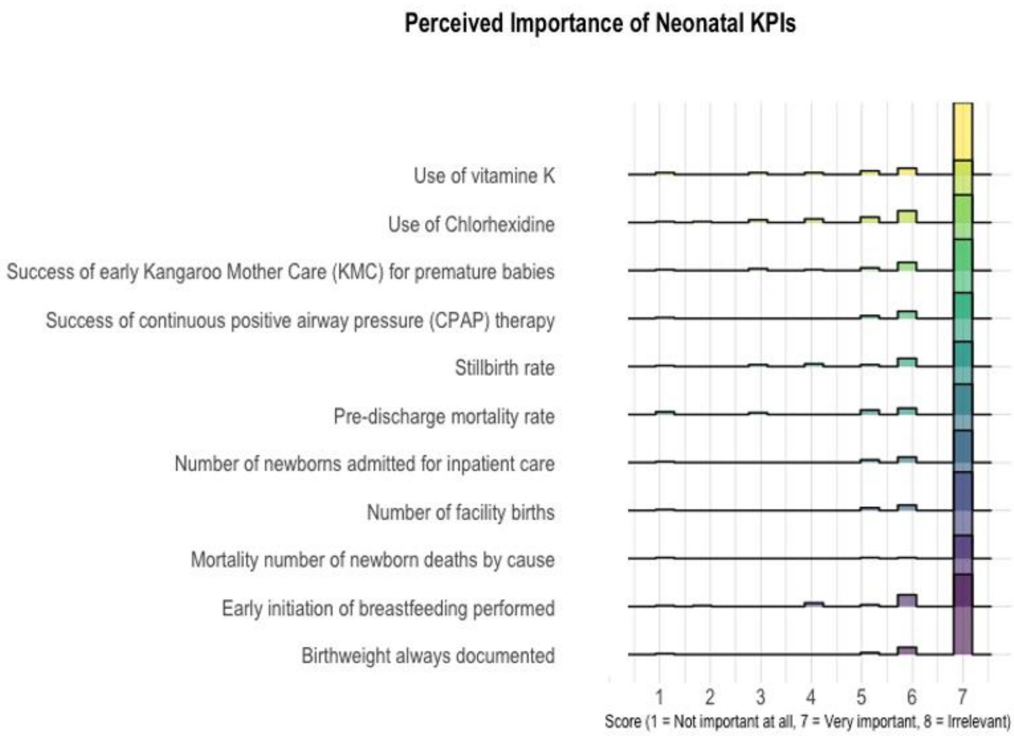


Figure 29: Ridgeline histogram plot for survey respondents' perception on the importance of neonatal KPIs to be presented in feedback tools, using a 1-7 Likert scale (not important at all) to 7 (very important), and 8 as the option to highlight a KPI as 'irrelevant'.

5.3.2.d. Main Users of Performance Feedback

This section identifies who participants believe are the main users of feedback dashboards with neonatal KPIs. By understanding the key stakeholders in the feedback process, we can better tailor feedback mechanisms to meet the needs of these primary users. Participants could choose with a required minimum of 1 and an optional maximum of 3 responses from a given list of 9 user group options (see Figure 30). Respondents perceived the Newborn Unit (NBU) Staff as the main users followed by Hospital Managers and Government Officials.

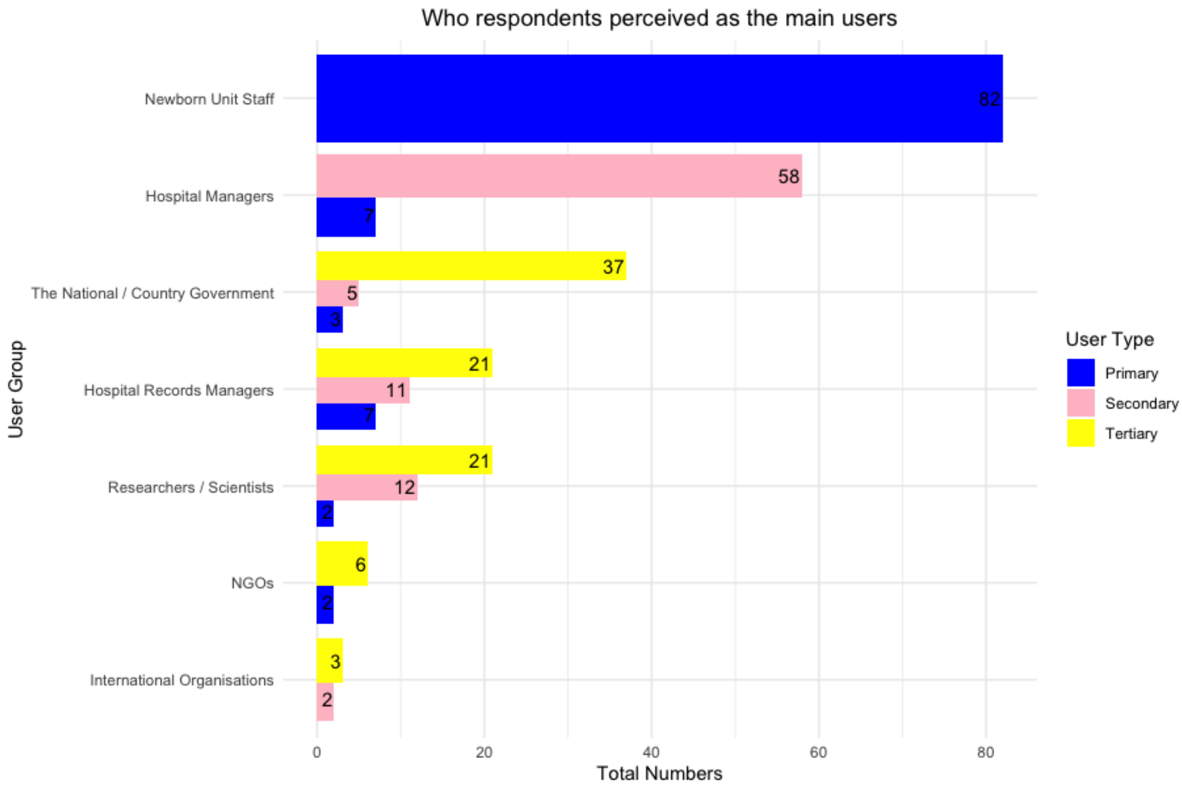


Figure 30: Users' perception on who the main users are of feedback on neonatal KPIs (content of the dashboard), with 3 multiple choice options of professions perceived as main user groups, presented as total responses. Each user group could receive a maximum of n=103 votes.

5.3.2.e. Preference of Feedback Delivery Frequency

In this part, I explore participants' preferences for the frequency of feedback delivery. This insight helps in designing feedback systems that align with user expectations and ensure timely and relevant information dissemination. The survey questions differentiated feedback on individual performance, team performance and organisation performance, and participants were given 8 response options: At any time/ real-time, daily, weekly, bi-weekly, monthly, quarterly, annually, and never.

For feedback on an individual level, respondents preferred receiving feedback at any time or in real-time (25.3%). 22.1% of the respondents would prefer monthly feedback and around

15% either daily, weekly or quarterly. For feedback on the team's performance, 40.2 % of the respondents preferred receiving feedback monthly, 26.8% weekly, 11.3% quarterly, 9.3% daily and 7.2% bi-weekly. Organisation or hospital performance was preferred by 46.8% every month, 33% quarterly, 7.4% real-time, 6.4% weekly (see Figure 31).

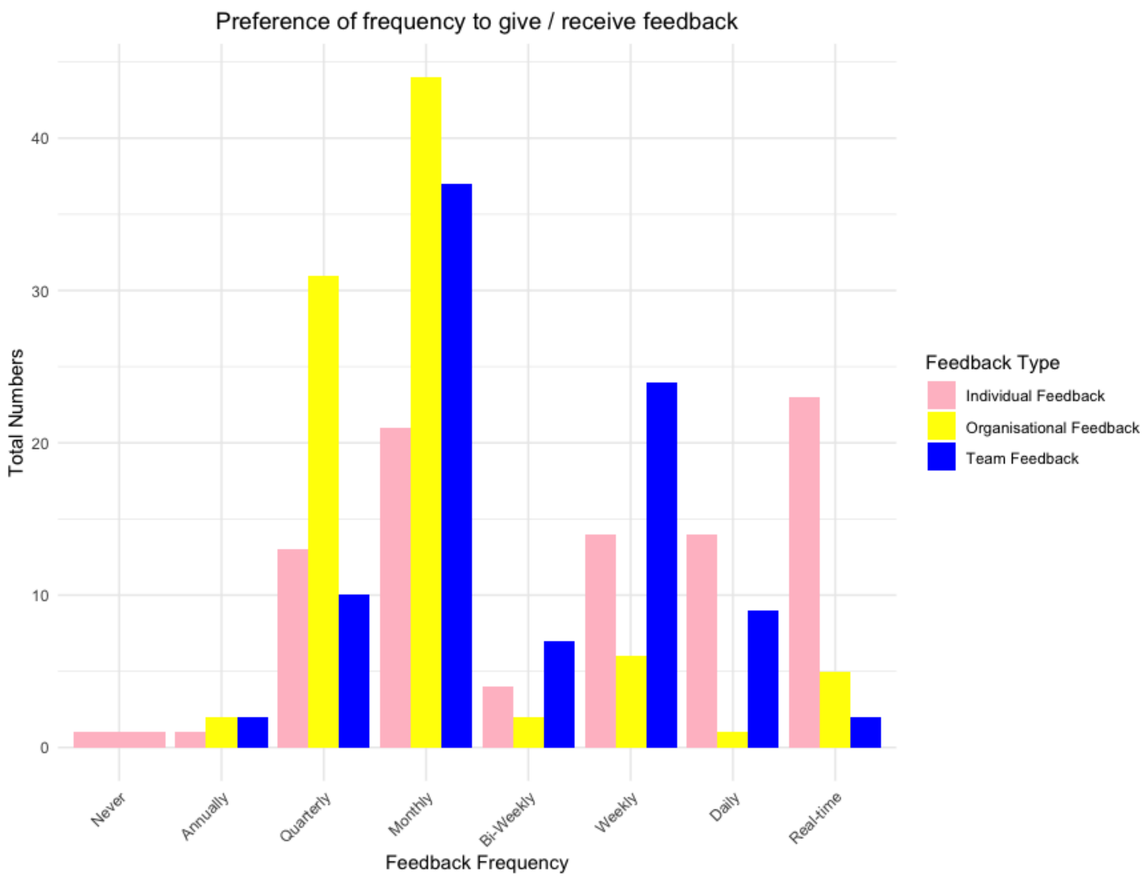


Figure 31: Users' preference on the frequency of giving or perceiving feedback on performance on an individual (blue), team (orange) or organisational (grey) level, with 8 multiple choice options for frequency on x-axis and total numbers of responses on y-axis.

Note: Numbers presented are totals. Each performance group (colour) could receive a maximum of n=103 votes.

The following simple majority can be drawn from the responses for participants having a preference in frequency of feedback: Preference for 1) individual feedback in real-time, 2) team feedback bi-weekly, and for 3) organisation feedback bi-weekly.

5.3.2.f. Confidence

Finally, this section presents participants' self-assessed confidence levels (as a cognitive self-measure) in improving QoC. By understanding their self-rated confidence, areas can be identified where additional support or training may be needed for performance improvement in neonatal care, and whether confidence influences respondents' ability to interpret information on the tool. Confidence of individuals and teams on their own knowledge and skills can also be an important component for cognitive test performance (as presented in Chapter II). Participants' confidence, wherein individuals had to assess their own skills, knowledge or ability was measured based on 5 different confidence scores using a Likert scale rating from 1 (not confident at all) to 7 (very confident). The 5 confidence points include the confidence:

a) to assess performance information on QoC.

b) how to use performance information to improve QoC.

c) to have the skills to improve QoC.

d) that I am able to achieve improvements in my hospital/ ward.

e) to know what feedback on performance and QoC means.

The correlation between the 5 different confidence scores is significant, with a correlation coefficient of 0.550 and higher. These five variables were subject to factor analysis using Principal Components Analysis (PCA), which determined that the five variables loaded onto one (new) variable, explaining more than 73% of the variance. The Kaiser Meyer Olkin test was .81, which determined that the sample size was sufficient for factor analysis. The Cronbach's Alpha was .91, suggesting a very reliable scale [17].As a result, the 5-item self-reported confidence measure was consolidated as one variable 'confidence', and calculated as the total (mean), for further analyses (see Figure 32*)*. All 5 groups showed a median of 6. Assuming participants responded with 6 (=confident) or 7 (=very confident), the majority of participants (n=74) showed high confidence in their ability to improve the QoC.

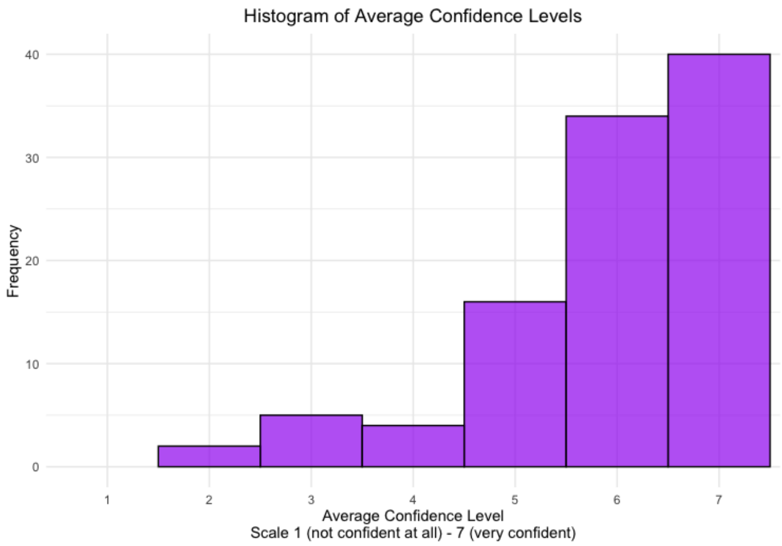


Figure 32: Users' (n=103) self-reported confidence as the sum of 5 confidence points, with 1-7 point Likert scale rating on x-axis using a scale 1 (not confident at all) to 7 (very confident) and total numbers of responses on y-axis. Note: Numbers presented are the mean of all totals.

i. User Perception of CIN and NEST Dashboard Design

This section explores users' perceptions of individual graphs after they had to interpret them with data taken from the responses to the survey tool questions for all Likert ratings from survey page 9 to 14 (Survey Outline, Appendix 5.8.), Self-reported perceptions of respondents were measured with the same 7-point Likert Scale to determine for each graph type (a) how easy they find the graph to interpret (Ease of Use) and (b) how useful they find the graph for feedback (Usefulness) on a Likert-scale between 1 (not easy at all) to 7 (very easy), and 1 (not useful at all) to 7 (very useful). In total, participants had to rate 6 graphs. For each tool, CIN and NEST, the 3 different graph types/parts were perceived by participants as relatively 'easy to use' and 'useful' as seen from the skewed distribution in Figure 33 below.

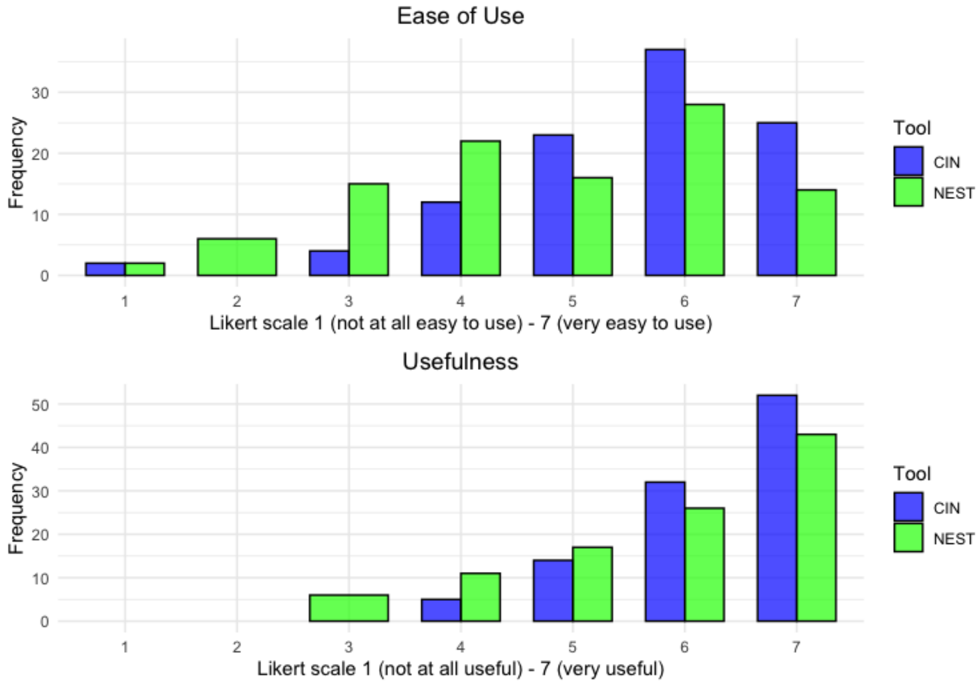


Figure 33: Users' perception on ease of use and usefulness (overall) after interpreting CIN and NEST graphs (mean), with 1-7 point Likert scale rating on x-axis and total numbers of responses on y-axis.

In addition, respondents perceived, on average, all CIN graphs a suitable way of giving feedback (Score Likert 1-7, average mean 5.984, average SD 1.216).

ii. Interpretation Success Rates of CIN and NEST Visualisations

This section focuses on the answers to sub-question 4 (see Table 5), which asked participants to solve tasks and interpret CIN and NEST dashboard feedback graphs as seen in Section 5.2.2.a. The data relates to the responses to the survey tool questions from survey pages 9 to 14 (Survey Outline, Appendix 5.8.).

All respondents (n=103) had to interpret visualisations from both dashboards as this task aimed to find differences between both tools, but in particular between different visualisation types. One graph (Appendix 5.3.) was excluded from the success rate calculations as the questions only provided perceptions rather than success scores. Consequently, the overall score for successful interpretation is based on 18 questions for 6 different visualisations (n=7 for 3 different CIN graphs, n=11 for 3 different parts of NEST). Respondents could select the correct response from multiple choice options, interpreting the graphs for important findings and the interpretation success rates were measured as correct (1) or false (0) for each question. Since the NEST dashboard shows three sections (top part, middle part and bottom part of the dashboard one pager) but the parts present different graph types, the results were measured based on the graph types rather than the three sections. The total interpretation success of both tools shows a normal distribution (Mean 0.569 SD 0.186), and overall, about 50% of all respondents could interpret the CIN (Mean 0.478/ SD 0.186) and NEST graphs (Mean 0.519/ SD 0.235) correctly (Figure 34). None of the participants could respond correctly to all questions, thus interpret all graph types correctly. However, two participants could respond correctly to all questions about the NEST visualisations.

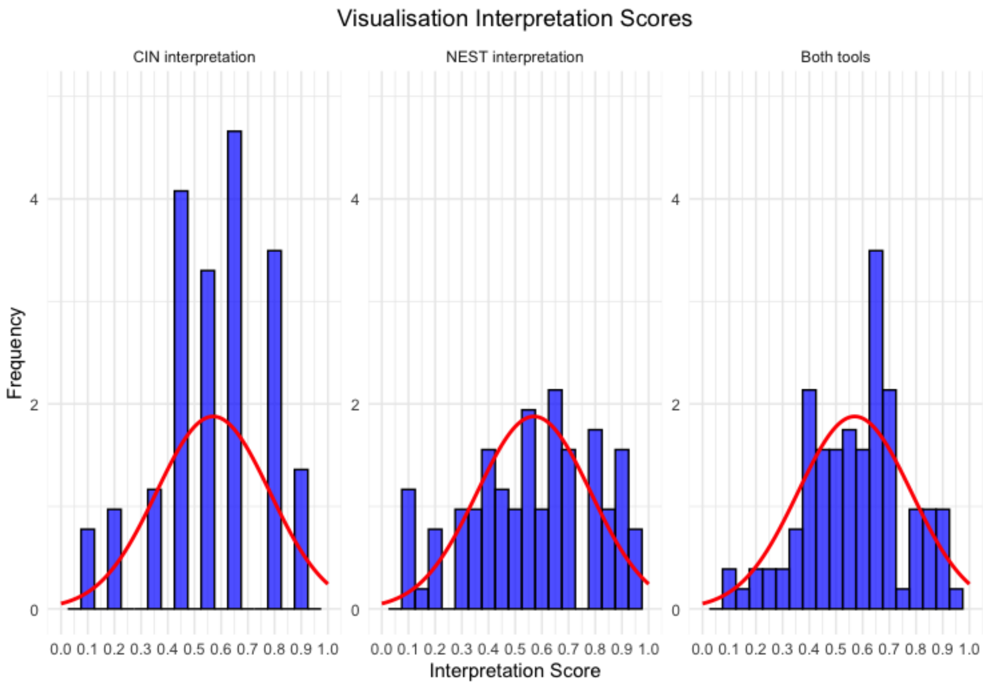


Figure 34: Histograms for interpretation success score of all respondents (n=103) for CIN and NEST interpretation success scores separately (based on 3 graph types each), and for total interpretation success score for both tools tested (based on 6 graph types) and using a test score of 0=false and 1=correct.

From the results of testing the interpretability of individual visualisations, a ranking identifies which specific graph types were more successfully interpreted than others, as presented in Table 8. The simple bar chart of CIN and the simple table of the NEST dashboard were interpreted more successfully than all other visualisations. Line charts for both tools were somewhat interpretable, with about 60% of participants responding correctly to related questions. While the complex table of CIN and the complex stacked bar chart of NEST received the least correct responses and, therefore, seems more difficult to be interpreted than other graphs.

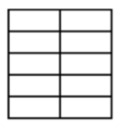
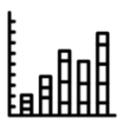
	CIN				**NEST**			
Rank	**Graph type**	**Correct responses**	**Ease of Use**	**Useful-ness**	**Graph type**	**Correct responses**	**Ease of Use**	**Useful-ness**
		in %	Median			in %	Median	
1	**Simple Bar Chart** **(1 question)** 	77.7 %	7	7	**Simple Table** **(3 questions)** 	73.4%	6	7
2	**Complex Line Chart** **(3 questions)** 	58.5%	5	6	**Simple Line Chart** **(2 questions)** 	59.2%	6	7
3	**Complex Table** **(3 questions)** 	34.3%	6	7	**Complex Stacked Bar Chart** **(6 questions)** 	38.5%	4	6
Total %	**All CIN graphs**	**50.9%**			**All NEST graphs**	**51.9%**		

Table 8: Ranking of individual graphs by interpretation success rate: Highest interpretation success (% of respondents interpreting correctly) compared to usefulness and ease of use perception score

responses by all participants (n=103) with a scale from 1 (not easy/not useful) – 7 (very easy/very useful). *Note: () Individual graphs were tested with a different number. For more than one question per graph type, the average success rate was calculated.*

The interpretation success rate results (column 3 for CIN and column 7 for NEST) are presented in line with users' perceptions of the 'ease of use' (column 4 for CIN and column 8 for NEST) and 'usefulness' (column 5 for CIN and column 9 for NEST) that were measured earlier (see Section 5.3.3.) with a 7-point Likert scale. Results show that users perceive all graphs in general as very useful and very easy to interpret, while the actual interpretability success rate varies, which indicates that users overestimate their ability to interpret the graphs. The association between the perception 'ease of use' and 'usefulness', and the ability to interpret, therefore, was tested next.

iii. Exploratory Bi- and Multivariate Analysis Results

This section focuses on the results to sub-question 5 (see Table 5), exploring associations between the success score to interpret different dashboard content and other pre-specified variables (relationships between results from Section 5.3.4 and 5.3.1.-5.3.3.). The statistical tests ANOVA, post-hoc test, Chi-square, and linear regression were used depending on the variables of interest. All other variables presented as part of the descriptive analysis (e.g., occupation in neonatal care) but not presented here did not show any association with interpretation scores.

5.3.5.a. Bi-variate: Association Demographics & Interpretation

Age, gender, profession and exposure to the tool have been selected to investigate these as factors of success in interpreting different graph types from the dashboard. Detailed statistical test results for each sub-sections are presented at the end of this section in Table 10.

i. Gender

Male participants reported using computers more often than females (r = -.299; p < .01) and scored slightly better on interpreting CIN graphs (M = .64; SD = .14) in comparison to the female participants (M = .53; SD = .22; t(93.526) = 3.185). Female participants in this data set were more likely to be holding clinical hospital roles compared with male participants (*chi square* (2) = 9.829; p < .01). There was no correlation found between gender and the interpretation success score for NEST or the total success scores.

ii. Age

Exploring age as a variable (six age groups as well as binary age groups younger/older participants, see Table 6), the analysis showed the older a participant, the less likely they are to use a computer (r = -.205; p < .05). However, there was no correlation between age and any of the interpretation success scores.

iii. Professional Groups

Professional groups were explored, as a pre-specified variable to potentially have an influence on the interpretation success scores. There were no correlations for individual professional groups and the success score of interpretability. When the three consolidated professional groups (Group 1 Researcher, Scientist & Others (n=12), Group 2 Information Professionals (n=36), and Group 3 Clinical Professionals (n=55)) were examined for an association with interpretation test scores, ANOVA tests showed there is a significant difference between groups (see Table 9). Data further suggest that Group 1 was more successful in interpreting graphs correctly than Group 2 and 3, but Group 3 was more successful in interpreting graphs correctly than Group 2.

Tools	**Graphs tested**	**Professional group**	**n**	**Mean**	**Std. Deviation**	$SE = SD \div \sqrt{n}$

	with n questions					
CIN	**3 graphs** tested with 7 questions	Researcher/ Scientist & Others	12	0.59	0.14	0.04
		Information Professionals	36	0.42	0.17	0.03
		Clinical Professionals	55	0.49	0.19	0.03
		Total	103	0.48	0.19	0.02
NEST	**3 graphs** tested with 11 questions	Researcher/ Scientist & Others	12	0.69	0.19	0.02
		Information Professionals	36	0.41	0.22	0.02
		Clinical Professionals	55	0.55	0.22	0.02
		Total	103	0.52	0.24	0.02
Both tools	**6 graphs** tested with 19 questions	Researcher/ Scientist & others	12	0.67	0.12	0.01
		Information Professionals	36	0.49	0.18	0.02
		Clinical Professionals	55	0.59	0.18	0.02
		Total	103	0.56	0.18	0.02

Table 9: Statistical results for success scores by professional group using a test score of 1=correct and 0=false. Interpretation performance of Group 1: Researcher/Scientist & others > Group 3: Clinical Professionals > Group 2: Information Professional.

In conclusion, the researcher group was the most successful in interpreting the graphs from both tools, followed by the clinical professional group, and then the information professional group. The distribution is presented in Figure 35. Therefore, the performance ranking is: 1. Researchers/Scientists, 2. Clinical Professionals, and 3. Information Professionals.

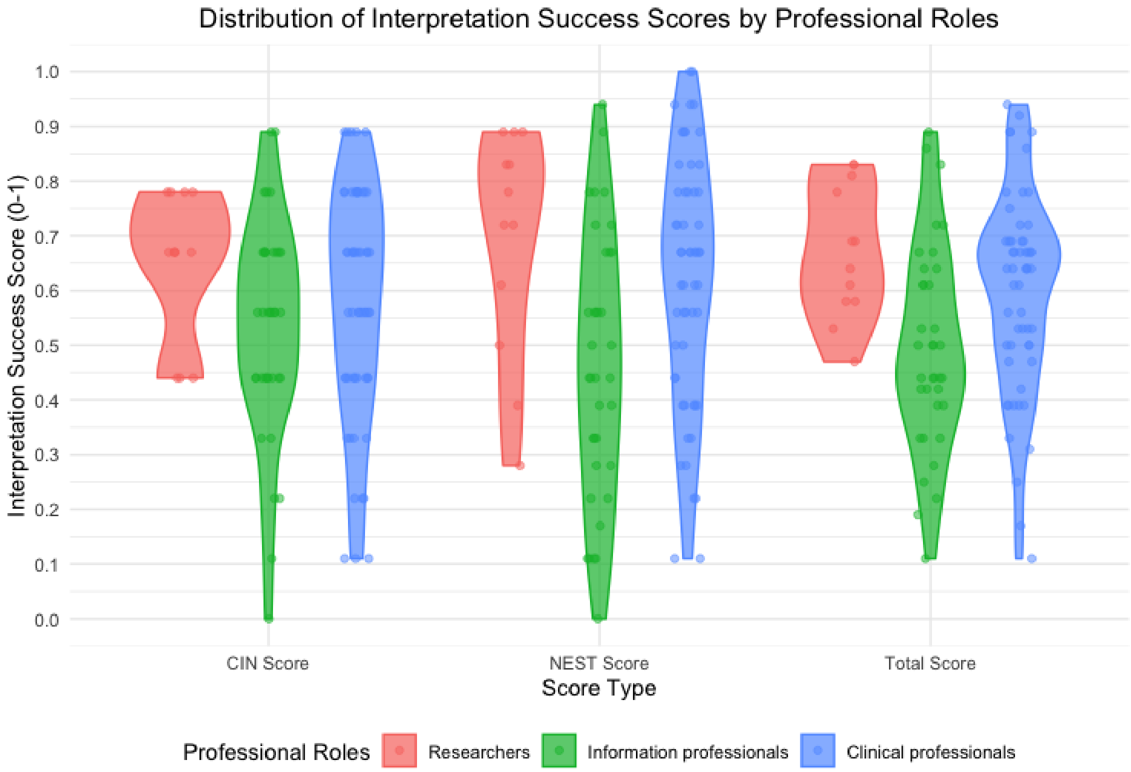


Figure 35: Violin plots to show distribution of interpretation success scores by professional roles using a test score of 1=correct and 0=false.

iv. Summary Table

Table 10 presents the association between various demographics and participant characteristics (gender, age, profession, and exposure to the tool) and the interpretation success scores for all graphs (Total), CIN graphs only, (CIN Score) and NEST graphs only (NEST Score).

Independent variable	**Dependent variable**	**Test**	**Test result**	**p-value**
Gender *1: Male* *2: Female*	**Total** interpretation	t-test	non-significant	
	NEST score		non-significant	
	CIN score		Male: M = .64; SD = .14; Female: M = .53; SD = .22; t(93.526) = 3.185	p < .01
Age *1: 21 to 40* *2: 41 or above*		t-test	non-significant	
Profession *1: Researcher/ Scientist* *2: Information professionals* *3: Clinical professionals*	**Total** interpretation	ANOVA, with post hoc	**Significant difference,** partial eta square = .112, a moderate effect size. F(2,100) = 6.413; Researchers (M = .83; SE = .14) Clinical professionals (M = .62; SE = .06) Information professionals (M = .33; SE = .08)	p < .01
	NEST score		non-significant	
	CIN score		non-significant	
Exposure *1: Exposed to one* *2: Exposed to two* *0: Not exposed*		ANOVA, with post hoc	non-significant	

Table 10: Summary table of statistical test results for association between i) demographics/ participant characteristics gender, age, profession and exposure to the tool (independent variables) and ii) interpretation success score of all graphs, CIN only and NEST only (dependent variables).

5.3.5.b. Bi-variate: Association Self-reported Perceptions & Interpretation

i. Training Needs

There was a correlation between the three training need questions/variables, so they were cumulated and used as one variable (see Section 5.3.2.b). Perceived training needs correlated negatively with the success scores in interpreting the CIN scores and, as a result, also with the total success score, meaning that the participants perception of their training needs correlates with their inability to interpret the graph types. Their perception of their needs to receive more training to make use of the information they receive with the dashboard, therefore, is legitimate.

ii. Confidence

There was a correlation between all five confidence variables; thus, confidence could be used as one variable (see Section 5.3.2.f). However, there was no correlation found between confidence and any of the success scores in interpreting the information given, meaning participants' own perceived confidence does not correlate with participants' ability to interpret successfully.

iii. Usefulness and Ease of Use

I found no correlation between the overall perceived usefulness and ease of use with any of the success scores in interpreting graphs. There is, however, a weak-moderate relationship between the perceived usefulness and perceived ease of use with the CIN line chart (Graph 3, Section 5.2.2.a).

iv. Summary Table

Table 11 below presents the association between user perceptions (training needs, confidence, ease of interpretation, and usefulness) and the interpretation success scores for all graphs (Total), CIN graphs only, (CIN Score) and NEST graphs only (NEST Score).

Table 11: Summary table of statistical test results for association between i) the user perceptions training needs, confidence, ease of interpretation and usefulness (independent variables) and ii) interpretation success score of all graphs, CIN only and NEST only (dependent variables).

Independent variable	Dependent variable	Test	Test result	p-value
Training needs *1: use the tool* *2: interpret content* *3: use feedback for change*	**Total** score	Spearman's rho correlation	Rho = -.251; The greater the training needs, the less ability to interpret TOTAL graphs	p < .05 negative correlation
	NEST score		non-significant	
	CIN score		Rho = -.208; The greater the training needs, the less ability to interpret CIN graphs	p < .05 negative correlation
Confidence *1: to assess performance information on quality of care (QoC)* *2: how to use performance information to improve QoC* *3: to have the skills to improve QoC* *4: able to achieve improvements in my hospital/ ward* *5: to know what feedback on performance and QoC means*	"	Spearman's rho correlation	Non-significant Participants' perceived confidence is not related to their ability to interpret graphs.	
Perception: Usefulness/ Ease of use *1: All graphs* *2: NEST graphs* *3: CIN graphs*	"	Spearman's rho correlation	Non-significant Participants' perceiving graphs as useful and easy to use is not related to their ability to interpret them.	
Perception: Usefulness/ ease of Use of Individual graphs 1: Usefulness 2: Ease of use	**Success score individual graphs:** **CIN:** Bar, Line, Table **NEST:** Bar, Line, Table	Spearman's rho correlation Pearson's rho correlation	Non-significant results except for **CIN line chart,** showing weak-moderate relationship 1: Rho = .371; 2: R = .210	1: p < .001 2: p < .05

5.3.5.c. Multivariate: Association Self-reported Perceptions & Interpretation

Several independent variables were then tested using the NEST, CIN, and total success scores (dependent variable) with multiple linear regression. All three scores were normally distributed (see Figure 34). 2 different models were tested, and a backward elimination conducted. Variables included in the overall model with non-significant results were:

- Gender
- Age
- Profession,
- Exposure,
- Training needs,
- Confidence,
- Perception Ease of Use,
- Perception Usefulness

The only significant results found were a negative correlation between training needs and success rate in interpretation of graphs, and the varied interpretation success by profession.

5.4. Chapter Summary

5.4.1. Further Research and Limitations

The quantitative component of the study faced challenges that have to be reported as limitations and should be considered when interpreting the results and guiding future research. Firstly, bias may be introduced by survey questions that influence the responses, for example, through the wording of the questions or the available answer choices [205]. However, the survey questions and answers were carefully selected with neutral wording, and I used the JISC software feature to randomise the order of response options and balanced rating scale response options (Likert 1-7). Measurement errors were minimised by following best practices from the literature for survey design [191]. Finally, I piloted and tested the survey to check any possible misunderstanding or influence and to delete unnecessary content.

Secondly, the survey tool JISC, recommended by the university for data protection, did not allow for time-tracking of individual responses, which can be important information in usability testing. Although I could observe participant reactions as I was present during the respondents' survey exercises, such as pauses during graph-related questions and extended time spent on graph interpretation, the exact time measurement data were not possible to collect. However, it was striking that participants spent most time trying to interpret the graph types with many participants, even with considerable time, unable to interpret these correctly, as my data show. Similarly, incorporating eye-tracking, which would have provided detailed insights into user attention and engagement, was not feasible due to the need for specialised expertise, additional ethical approvals, and increased complexity. By tracking eye movements to create visualisations like gaze plots and heatmaps one can significantly enhance understanding of user interaction with a dashboard, allowing designers to optimise layouts for improved usability and user experience [206]. Since both dashboards were in the prototype phase and not entirely web-based at the point of this study, I focused on the interpretability, ease of use and usefulness of the content rather than testing systems usability, recommending to add time-and eye-tracking testing for when the dashboards are digitally more developed.

Lastly, the sample size could be criticised. However, the study is exploratory, focusing on the unique contexts of 20 CIN and the sub-set of 13 NEST hospitals with approximately 20 dashboard users for each facility (see Section 3.4.). The number of respondents exposed to CIN and NEST dashboards is therefore naturally limited. The mixed methods design allows for triangulation and a larger qualitative data set enhanced the study's robustness despite a modest quantitative sample and these data will be presented next. Ethical and logistical constraints due to the COVID-19 pandemic also restricted data collection as priorities in that period were to refrain from interrupting clinical work if possible. The high response rate and focus on user experience and usability testing for a specific user group (as described in Section 5.2.2.b) justifies smaller sample sizes in similar studies. Finally, the sample size of 103 participants was sufficient to detect meaningful effects in the context of this study. The survey provides valuable insights, serving as a foundation for future dashboard design iterations and research in this area.

5.4.2. Dashboard Performance: What People Say and Can Do

The quantitative assessment was an important part testing what users can do with dashboards as they are currently designed; thus, it helps to reflect on the reality of how interpretable they are, and therefore, whether data is likely to be usable to improve QoC. For the user experience empathy map the survey provided information on what users can do and reflections on their perceptions (see Figure 36).

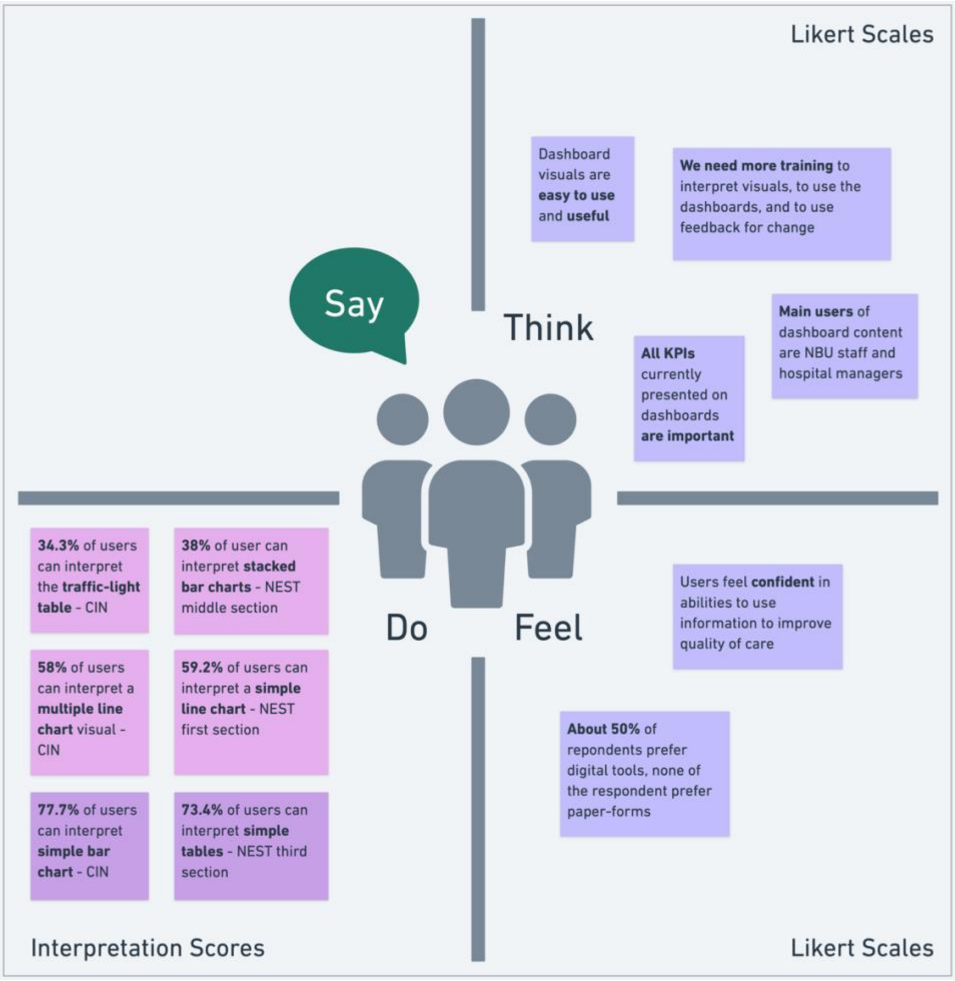


Figure 36: User empathy mapping exercise using quantitative data - What people can ‘do’ differs from what they think they can. Created by CH, inspired by Empathy Mapping by NN Group [89].

The quantitative assessment on what dashboard users can interpret and what their perceived needs and preferences are, highlighted that the current dashboard designs are not as ‘easy to use’ and ‘useful’ as users say they are. The findings provided evidence on what parts of the dashboards users struggle to take information from, offering guidance and which parts need a re-design. Table 12 below presents the main findings of the survey as a summary. In the next Chapter VI, I will report the findings from the qualitative data, which will add more insights into what users say, think and feel

Table 12: Summary of findings from the survey presenting 3 key findings, other observations, positive highlights and additional comments of participants.

Metric	Description	Results
Main objective	The survey mainly aimed to explore (a) user experience more broadly, and (b) usability in terms of current dashboard designs.	The objectives were tested with (a) Likert-scales and multiple choice questions, and (b) interpretation success rates of dashboard graphs.
Total respondents	The total number of people who completed the survey.	The survey included 103 participants from the Kenyan healthcare setting.
Response rate	The percentage of people who responded to the survey out of those who were asked.	100% response rate and no drop outs.
Key finding 1 ***Interpretation success***	Overall interpretation success rates.	About 50% of participants cannot interpret the information. No participant could interpret all or none. One participant could not interpret any of the graphs for CIN. Two participants could interpret all graphs/ respond to questions for NEST.
Key finding 2 ***Non-affecting influences***	Factors that are not influencing interpretation success rate.	Age, gender, exposure, and users' perceptions and confidence show no (or only little) effect on interpretation success rates.
Key finding 3 ***Affecting influences***	Factors that are influencing interpretation success rate.	**Profession:** The researcher group shows higher success rates than clinical professionals, and the latter show higher success rates in comparison to information professionals to interpret dashboard graphs. **Training needs:** There is a negative correlation between respondents' perceived training needs and interpretation success rate, meaning those who said they have training needs were less successful in interpreting graphs correctly.
Other observations	Observed response time and breaks when completing the survey.	Participants seemed to respond quickly to survey questions but paused for longer periods of time for the questions to interpret of graphs.
Positive highlights	Respondents' confidence to improve care for newborns.	Although not affecting interpretation, participants are confident in terms of being able to improve QoC, rather than being demotivated.
Additional comments	Respondents were given the opportunity to give feedback on the survey at the end.	Respondents highlighted mainly the importance of this study and that they want to be more involved and/or asked about their user experience.

Chapter VI: Mixed Methods – Qualitative Inquiry: Implementer and User Experiences

In this Chapter, I present the results of the qualitative component of the empirical study. After a quick recap of the objectives, I present how the qualitative data adds important information to the quantitative component, including the benefits for interviewing dashboard implementers and users. I provide more details about how the literature informed the interview tools and the study. I follow a structure to present the findings based on whether they are people-, tool- or environment-related. Finally, I summarise the findings before integrating both the quantitative and qualitative components in Chapter VII

6.1. Introduction

6.1.1. Recap of Objectives

The qualitative component of this mixed methods study aimed to answer research questions 2 and 3, as outlined in Chapter III, to explore the experiences, perceptions and values of implementers and users, contextual influences and the implications moving forward. It adds to the quantitative findings to contribute to the development a systematic evaluation and iterative re-design strategy for future dashboards. I present findings from 42 in-depth, semi-structured interviews collected as two data sets (see Figure 37), which offer a deeper dive into implementer and user experiences about and perceptions of the CIN and NEST dashboards.

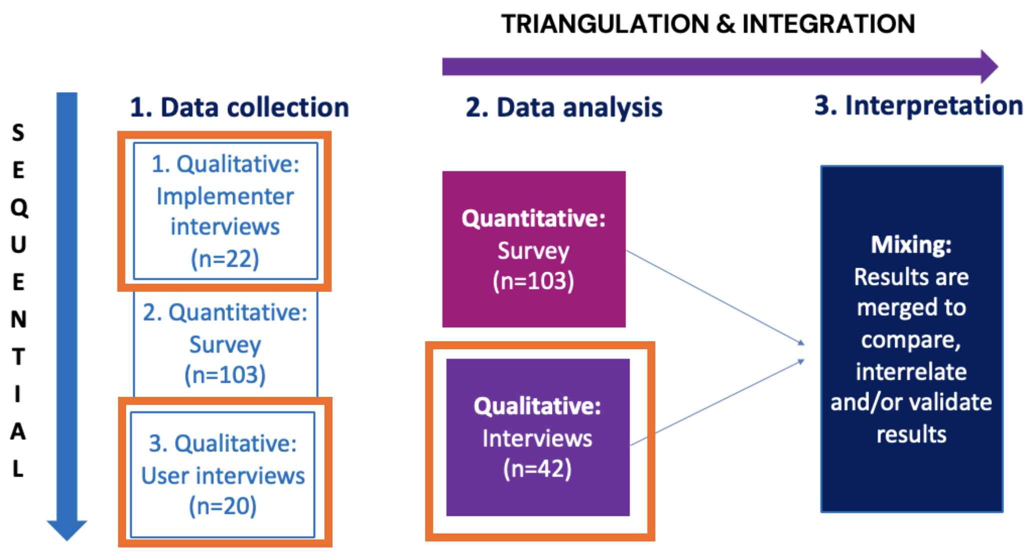


Figure 37: Overview of mixed methods data sets – Qualitative component

6.1.2. Aims and Contribution of Qualitative Data

6.1.2.a. Two Perspectives in Qualitative Component

My study investigates dashboards as the (digital) medium for delivering data-driven feedback on performance. The qualitative component helps to understand how people interact with dashboards and thus helps to improve these interactions and experiences. Here, people are split into those who designed and implemented the tools (the **implementers**) and those who are targeted by them to use the tools (the **users**). By interviewing both, we can investigate the dashboards from two angles (see Figure 38 below), since people may interact with dashboards differently. By interviewing tool users, it is possible to explore how and why people use the tools (or not), why some visualisations are interpreted more successfully than others, and identify the challenges involved and how to overcome them. The qualitative data also explores how the perceptions of the implementers and users overlap or differ. Implementers' views add to our understanding of why and how they made decisions on feedback tool design(s), the KPIs and implementation processes, and what they already tested during the development process.

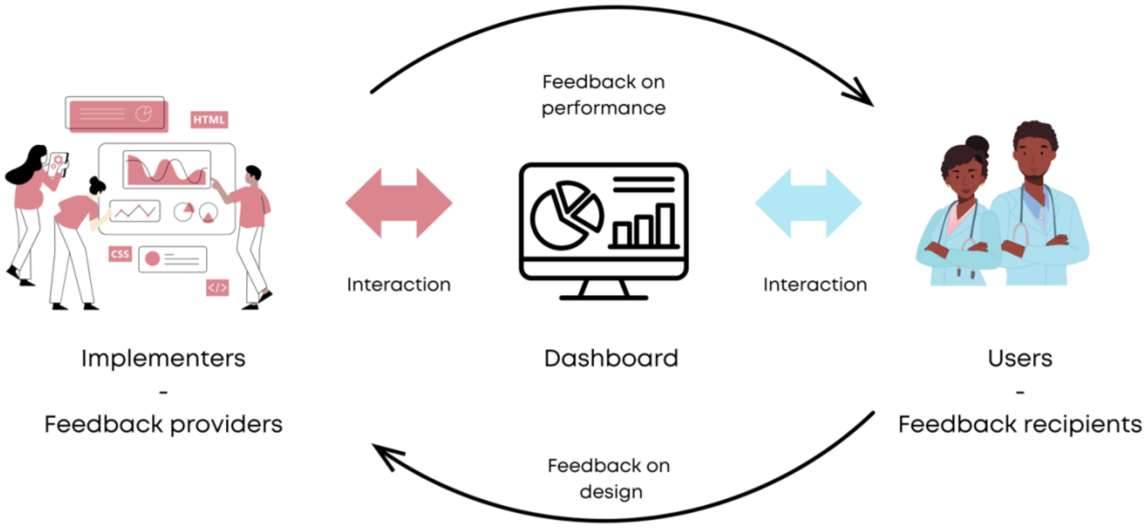


Figure 38: Two perspectives to understand dashboards to provide feedback on clinical KPIs.

6.1.2.b. Qualitative Data Complementing Quantitative Component

The quantitative data revealed some specific usability issues, showing whether dashboard users can or cannot read correctly from the current dashboard feedback design. The interpretation success scores indicated only modest success. Likert-Scores showed that participants perceived the dashboard mostly as very ‘easy to use’ and ‘useful’. However, the same individuals often could not interpret the graphs in the dashboard. A purely quantitative assessment cannot capture the wider user experience about how the tool is useful and why it is easy to use (or not). Thus, this chapter adds depth and richness to the quantitative findings of the binary result of whether people can or cannot interpret by adding insights into why users struggled to interpret dashboard information, alongside implementers’ experiences and knowledge they applied to develop the dashboards. In Table 13 I show how the collected data are linked to the examination of usability and broader user experiences, and how the qualitative complement the quantitative data. This draws on the design and A&F literature, particularly on CP-FIT (see Figure 3).

Table 13: Broad overview of how dashboard user experience has been tested

	Investigation elements/ tests	Metric	Data	CP-FIT cycle **Number in () relate to the CP-Fit feedback cycles stages.**
WHAT PEOPLE CAN DO	**USABILITY** measured relative to user's performance on given set of test tasks in survey: How many people can interpret how much.	**Success rate (error rate):** The percentage of users who can complete a specific task (and percentage of users who can't/ encounter error)	**Quantitative survey data** providing behavioural information	Testing whether feedback (3.) allows participants to improve their clinical performance (10.) and change behaviour (9.)
		Time to complete task: Average time users need to complete task as a reflection on the difficulty of the task	**Observational notes** without actual time tracking providing additional behavioural information alongside success rate.	Observing participants' interaction (8.) and behaviour (9.)
WHAT PEOPLE THINK, SAY, FEEL	**USER EXPERIENCE** measured relative to user's perceptions: Why & how can we fix it	**User's subjective satisfaction score:** How users perceive their own satisfaction with the tool.	**Quantitative survey data** (Likert)	User experiences as perceptions (5.) particularly on feedback (3.) but also on acceptance (7.) and intention (8.)
		User's subjective satisfaction description: How users perceive their own satisfaction with the tool.	**Qualitative interview data** providing more attitudinal information, people's perceptions, thoughts, beliefs,	User experiences as perceptions (5.) particularly on feedback (3.) but with

		Interest: How interested users are to use the specific tool (vs. others).	self-reported needs and preferences *Note: Attitudes can be interpreted as behavioural when becoming a behavioural pattern or practices. Attitudes can positively or negatively affect a person's behaviour*	reflections on all other parts of the feedback cycle (1.-11.)
		Barriers: What users perceive as barriers to use the tool.		
		Efforts: What users perceive in terms of how much effort it requires to use or trying to use the tool.		
		Needs and preferences: What users think they need or prefer (like, don't like).		
		Suggestions: Users perceptions of what else could be better and/or providing a broader vision of what action points are needed.		
	IMPLEMENTER EXPERIENCE measured relative to implementer's perceptions	**With the above considered, examine what else they can report on, what worked well and what didn't**	**Qualitative interview data** adding information from the implementer perspective	Implementer experiences as perceptions (5.) around the whole feedback cycle (1.-11.)

6.1.2.c. Exploring CP-FIT and Design Elements

In conducting this empirical study, I explored the 3 CP-FIT variables that influence feedback and elements of the CP-FIT cycle from stage 1 to 11 (see Figure 3 and Table 13).

In Table 14, I illustrate how the interviews provide insights into three overarching domains. These domains are partially informed by the CP-FIT framework, which Brown et al. describe as three variables influencing the feedback cycle: recipient variables, feedback variables, and context variables. While I draw connections to these CP-FIT variables and the framework as guidance to evaluate feedback interventions, I incorporate other relevant literature as outlined earlier, ensuring that the findings are not restricted to CP-FIT. Therefore, I report findings structured around three domains to summarise the qualitative findings:

(1) **People:** This domain relates partly to CP-FIT's recipient variables, focusing on the roles and characteristics of those receiving feedback, but I extended it to include feedback implementers.

(2) **Tools:** This domain relates partly to CP-FIT's feedback variables and links to the feedback mechanisms themselves. I extended this part by looking into dashboard specific details as the medium to provide feedback and added design principles as outlined earlier.

(3) **Systems:** This domain corresponds to CP-FIT's context variables, considering the broader contextual factors affecting the feedback cycle. I extended this to the larger, interconnected structures and processes of health systems, the specifics of health information systems into which dashboards ideally integrate.

Table 14: Information that user and implementer interviews provide for each domain

DOMAIN	USER INTERVIEWS, adding information…	IMPLEMENTER INTERVIEWS, adding information…
PEOPLE 〉 *Links to CP-FIT Recipient variables*	…about the role of individual users or user groups, their perceptions on role specific factors, e.g., data literacy, training needs, teamwork.	…about what implementers learnt from engaging with individual users or user groups, but also from each other as within the implementer team and other research groups, but also how they reflect on their own role(s).
TOOL 〉 *Links to CP-FIT Feedback variables*	…by providing a deeper tool specific understanding, the usability, ease of use and usefulness of feedback tools e.g., the design, colours, benchmarking etc.	…how implementers designed feedback tools, their intentions and values of the tool.
SYSTEM 〉 *Links to CP-FIT Context variables*	…on the barriers and contextual influences within the existing health and information systems where the tool is implemented, e.g., interoperability, workload, staffing, power cuts.	…how implementers implemented feedback tools and integrated into existing system(s), what the challenges or unexpected barriers or consequences have been.

6.2. Methods

Having introduced the rationale for the qualitative work in Section 6.1., I describe here the methods used for the qualitative component, building on the methodology in Chapter III.

6.2.1. In-Depth, Semi-Structured Interviews

In-depth interviews were conducted with implementers using an interview guide to explore how the tools were designed and implemented and their expectations of how these tools would help users in their daily practice (Implementer Interview Tool, Appendix 6.1.). Interviews with users focussed more on the experience and perceptions of using the feedback tools deployed by CIN and NEST and wider contextual experiences affecting the usability of the tools (User Interview Tool, Appendix 6.2.). Open-ended questions and probes were informed by the a) A&F and design literature (Chapter II) and b) scoping review results (Chapter IV).) and evolved over the interview process to enable specific areas or emerging interesting themes to be explored.

A user journey mapping exercise was initially considered to outline how the dashboard feedback integrates into users' daily practices (each CP-FIT cycle step) and learn more about their experiences using the CIN and NEST tools in their daily routine. This was planned to take the form of a 'walk-through' interview to ask questions during people's day-to-day work at the facilities and observe their dashboard use within the facilities. However, due to COVID-related travel restrictions and limitations on visiting hospitals in Kenya, the pragmatic decision was made to collect this information as part of more standard interviews as a cognitive, narrative 'talk-through' of the tool interface and information instead.

6.2.2. Data Collection

In total, 42 participants were interviewed, of whom I considered 22 implementers and 20 users. Two of the implementers expressed having a double role, as user and implementer, due to their employment at one of the partner hospitals.

6.2.2.a. Sampling

As qualitative research follows a theoretical rather than statistical logic, I employed a sampling procedure focused on the 'purposive' selection of groups or subsets of the populations of interest [207]. In purposive sampling, potential sources of rich and contextually relevant information are identified to obtain an in-depth understanding of the specific phenomenon or area of interest [208]. Applying these principles to this study, participants were selected expected to provide the most insightful data about the use and impact of the CIN and NEST dashboards in order to inform the design and implementation of more effective dashboard designs and implementation strategies in the future. I attempted to interview my survey participants, but data privacy restrictions prevented follow-ups with everyone. Finally, all potential dashboard users were considered suitable participants. Despite this, 10 out of 20 user interviewees (50%) had also completed the survey. Based on a pragmatic decision during the COVID-19 pandemic and limited research funding, I only interviewed users in Kenya, with a purposeful selection that covers different professional types and dashboard exposure.

6.2.2.b. Participants

The study included 22 implementer interviewees from CIN and NEST with different roles within the programme, such as programme and data managers, clinicians, researchers, QI team on the ground and facilitators. 9 were female and 13 were male. The 9 NEST implementers had some transnational roles due to the programme focusing on multiple countries, specifically Kenya, Malawi, Tanzania, and Nigeria. The 9 CIN implementers and the 4 who had experience with both CIN and NEST, primarily focused on implementation in Kenya. The study further involved 20 user interviewees, comprising 6 male and 14 female participants, including nurses, paediatricians, health records information officers (HRIOs), and data clerks. 10 participants were exposed to both tools, CIN and NEST, and 10 only to CIN (with no prior exposure to the NEST dashboard). For detailed participant information, see Appendix 6.3. for implementers and Appendix 6.4. for users.

6.2.2.c. Interview Process

Due to pandemic-related travel restrictions, I conducted all implementer interviews online and worked with a local research assistant (RA) to conduct all user interviews in person. Data collection was performed through a continuous, iterative process that involved regular sense-checking and ensuring the interview data covered the relevant domains and metrics of interest outlined in Table 13 and 14. Interviews continued until thematic saturation was felt to have been achieved. For both data sets, 3 pilot interviews were conducted to adjust the interview tools with minor changes (e.g., order of questions or explanations) to improve the interview flow.

The user interviews were all conducted in person. I conducted the pilot interviews with the RA before the RA continued to interview more users at their facilities. After every 2-3 interviews, we reviewed the interview transcripts and notes to see whether we needed to make any adjustments. Data saturation was felt to have been reached after approximately 12 interviews for both data sets. However, another set of participants was interviewed to ensure sufficient representation of each professional group, including tool users and implementers (CIN and NEST). As a result, at least 4 people per user group were interviewed, observations collected, and content reviewed regularly.

6.2.3. Data Management and Analysis

Data was collected using audio recording, with the implementer interviews recorded through MS Teams and the user interviews through a manual, digital recorder, anonymised and transcribed verbatim, and deleted from the recorder afterwards. The transcription was outsourced to a professional transcription service provided through KEMRI. All recordings and transcriptions were password protected and transferred using online secure platforms such as OneDrive, in accordance with the University of Oxford and KEMRI guidelines. I checked and corrected the audio and transcripts several times, which also helped me immerse myself in the data from all the interviews. Interview dynamics were captured in the form of fieldnotes, to complement findings if necessary.

I applied a modified framework approach to analyse the data, starting with a coding framework that drew on elements of CP-FIT and the design literature (Chapter II). I used MAXQDA by Verbi as a data management programme [209]. The framework method was used to systematically analyse the data, allowing for the identification of key themes and patterns related to users' experiences, needs, preferences, and barriers in interpreting dashboard information [126], [210]. Framework analysis is a qualitative method known to be suitable for applied research where a policy or practice is of interest [126], [211]. The approach to analyse the data involved five steps, divided into a first phase of summarising and making sense of the data (familiarisation and developing a coding framework), and a second of interpreting the data (charting and analysing the data based on themes):

Phase 1: Summarising and making sense of data

〉Stage 1: Familiarisation

〉Stage 2: Developing a coding framework

〉Stage 3: Coding or indexing

Phase 2: Interpreting data

〉Stage 4: Charting

〉Stage 5: Mapping and interpretation

Phase 1

I developed one coding framework for each set of interviews, with some overlaps between them. In both cases, the frameworks were based on a combination of the interview guide (based on the literature presented earlier and Table 13 and 14), and a close reading of transcriptions. I continued to add new themes as I coded. In this way, the coding process was primarily deductive to ensure key areas of interest were examined, but also had an inductive element to allow for emerging findings and unexpected ideas and experiences to be captured. The first framework (implementer interviews) was very detailed, whereas the second (user interviews) retained the main themes and, thus, collapsed and clarified the sub-themes to reduce unnecessary overlaps. Therefore, the final coding frameworks (Appendix 6.5.) were developed through the analytical process.

Phase 2

I began the analysis by charting the detailed coding findings and linking them back to CP-FIT and design domains from the literature. My charts differentiated between data on CIN and NEST, and between inputs from different professional user groups. The charts and analyses were structured to cover main areas of interest like dashboard design, tool implementation, participant engagement, and system integration. Tables 13 and 14 laid the foundation for these themes for both data sets. This structured approach ensured that the analysis was thorough and reflective of the diverse perspectives and insights of participants captured through the interviews while also allowing me to examine these alongside the themes identified from the literature and add additional details. All coding themes and variables of interest are displayed in Table 15.

Table 15: The 5 coding themes, their sub-themes, how they link to CP-FIT, and how they are arranged in the three domains and chapter themes 'people', 'tool' and 'system'.

5 Coding themes	Sub-themes	CP-FIT themes	Domains
1. Key roles and people involved	• Interviewee details • Primary and secondary user • Other contributors • Skills • Knowledge • Workload	**Recipient variables** • Health Professional characteristics • Behavioural response	**PEOPLE**
2. Designing a dashboard	• Design principles or guidelines • KPIs • Usability (Easy to use/ useful) • Graph types for translating data into information • Benchmarking • Data sources, collection and quality • User engagement/ Co-design	**Feedback variables** influencing the cycle **& feedback cycle steps**	**TOOL**
3. Implementing the tool	• Accessibility • Delivery • Training • Response mechanisms for users to give feedback • Integration existing HIS • Evaluating design	**Context variables** (tool specific implementation process) influencing the cycle **& feedback cycle steps**	**TOOL**
4. Engaging with the wider system and contextual factors	Other contextual factors that are not directly linked to tool design and implementation but have an effect, e.g., • Staff shortage • Power cuts • Workload • HIS challenges	**Context variables** • Organisation or team characteristics • Co-interventions	**SYSTEM**
5. Purpose and impact of the tool	• Information needs • User target groups • Expected vs actual behaviour change • Specific tasks the tools help to achieve • Future potential	**All feedback cycle variables**	**PEOPLE** **TOOL** **SYSTEM**

As mentioned in Chapter III, I used two checklists to assess the qualitative component of the study: The Critical Appraisal Skills Programme (CASP) was used to ensure methodological rigour, validity, and relevance [138] (Appendix 6.6.). In addition, I used the Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist to support the appropriate reporting of results [137] (Appendix 6.7.).

6.3. Findings from Qualitative Data

Here, I will present the main findings from the 5 main coding themes by aligning them with the three overarching domains outlined above. The coding themes fit into the three overarching domains: 'Key roles and people involved' falls under **People** (6.3.1.), the two themes 'Designing and implementing the dashboard' under **Tools** (6.3.2.), and 'Engaging with the wider system and contextual factors' fits within the **Systems** (6.3.3.) domain. The theme 'Purpose and impact of the tool' intersects with all three domains, reflecting its relevance across different parts of the CP-FIT feedback cycle. Within each section, I report what respondents told me that fits into these domains as factors with potential influences on the value of the dashboard and its aim to improve QoC for newborns (see Table 15). In this way, I differentiate interviewees' different roles and their perceptions, and the effects of wider contextual factors. Interview quotes if not added within the text can be found in separate tables for each sub-section in Appendix 6.8 - 6.10.

6.3.1. People Involved with Dashboards and their Roles

This section presents findings about the interviewees, including their roles, who they describe as developers and users of the dashboards, and other people involved. This background information helps to comprehend how interactions with the tools occur within the hospital context and on the implementer side. As outlined in Chapter II, identifying primary and secondary users, and adding their knowledge and experiences can be important for ensuring inclusivity of different needs.

6.3.1.a. Implementer Roles

As noted above, 22 dashboard implementers were interviewed across both programmes: 10 from CIN, 9 from NEST, and 3 with roles in both CIN and NEST. The interviewees had varied professional backgrounds, including clinical work, research, study coordination, data and information management, and quality improvement. At the time of the interviews, CIN implementers were employed by KEMRI, while NEST implementers were contracted through the NEST programme. Some participants had additional affiliations with the University of Oxford, the London School of Hygiene and Tropical Medicine (LSHTM), Rice University, or partner hospitals.

The **CIN implementer** team is reported to be a relatively large team, consisting of clinicians, study coordinators, research officers, data managers, data assistants and statisticians, with multiple roles in the dashboard cycle from data collection to delivering feedback to users. Besides using the tools to track changes over time and identify opportunities for improvement, the CIN feedback implementers are predominantly involved in data collection, management, analysis and quality assurance (Appendix 6.8., Quote 1, 2, 3). In their description of what their role is, it is evident that they perceive themselves as the designers and implementers of the CIN tools. They are the "*feedback providers*" and the "*tool designers"* with the Principal Investigators (PIs) together with their research team making the decisions on how dashboard feedback is designed and delivered, and ensuring the users understand it by listening to their response (Appendix 6.8., Quote 2, 4, 5).

NEST implementers, with a similar team in terms of their background, e.g., clinicians, statistical and data personnel, and programme managers, reported more about how they are involved with ensuring the NEST dashboard is used to support the quality improvement efforts progressing on the ground (Appendix 6.8., Quote 6). As the NEST programme is implemented in Kenya, Malawi, Nigeria and Tanzania, they also have country team lead personnel, including information systems leads and data collectors on the ground (employed by the government) and a county-level quality improvement team (Appendix 6.8., Quote 7). Like CIN, the NEST implementers decided on the dashboard design and which KPIs were included, with decisions influenced by feedback from individual users (Appendix 6.8, Quote 8).

6.3.1.b. User Roles

20 dashboard users with different hospital roles, such as paediatricians, nurses, healthcare HRIOs and data clerks with varying experience levels, were interviewed (Section 6.2.2.b). When responding to questions about demographics, roles, and tasks, users indicated they feel responsible for improving care and data quality, interacting with dashboards and feedback systems in various ways.

Paediatricians see their primary role in clinical care, teaching, supervision, and participation in research. They see themselves as the people with the main responsibility for outcomes and QoC improvements. For example, one paediatrician reported that they called a colleague from a different facility to discuss numbers from a feedback report and learn from them what they (can) do to improve outcomes when faced with similar limited resources and staff shortages.
Nurses' self-perceived role is to deliver QoC for neonates, aiming to reduce the duration of hospital stays and prevent infections within the unit. They also advocate for the babies, facilitating communication between healthcare workers and caregivers. The (deputy) nurse in charge is responsible for creating a fair working environment by ensuring adequate supplies and managing staff effectively. They also focus on distributing policy information, supervising staff performance and teaching students. Furthermore, they say that they educate and update staff on new guidelines and technologies, ensuring the team is equipped to handle the latest changes in neonatal care. Nurses perceive their role as essential not only to the care of newborns but also to the broader objectives of QoC and the health system as they are

committed to continuous improvement and patient care. The dashboards support them "*having the status quo*" of their unit's performance providing "*a snapshot*" of conditions and outcomes alongside these efforts. Data clerks' self-perceived role is to use data to produce reports and differs from the clinical roles. They use language that distinguishes between *"they"* (referring to clinicians, including doctors and nurses) and *"us"* (the data people).

- *"…my active role is first [the] production of the report and just making sure it's the correct report, counterchecking and also presenting the same report." – Data clerk Interview 4 Line 40*
- *"Mostly [in] the hospital when it comes to the reporting, they use the CIN report because the one they do [routine hospital data] ...will not have captured everything. I think they miss [capturing everything] while doing maybe the admission, the discharges, they miss a lot of things. Sometimes they don't record them in the admission books, and they don't discharge when the referrals are not recorded. But me, I use the physical file so I'm able to capture all the records." – Data clerk Interview 7 Line 47*

The second quote above suggests that data clerks also distinguish their roles from those of the HRIOs. The reason could be that HRIOs are official hospital employees, while clerks usually comprise members of the CIN/NEST project team within the hospital. However, HRIOs and data clerks both perceived their tasks mainly in: Data entry, data and error cleaning, synchronisation of the data, maintenance of confidentiality (records and data), filing records after data entry, retrieving data for data quality assessment, data-based generation of a report, counter checking reports and presenting results (which can include to visualise data), ensuring that reports are complete, accurate, timely, and transferring the paperwork to computer and electronic systems, e.g., DHIS2 and Redcap for CIN/NEST. The goals in this role are reported mainly as (a) data collection and utilisation, (b) evaluation and monitoring of data systems and indicators, and (c) improving data quality and quality of hospital care. HRIOs recognise the importance of their own role to ensure data is collected and utilised by others (managers and clinicians) so that data-driven decisions are made to improve quality of hospital care, for example one saying:

- *“So, when we're doing the improvement of the hospital, we rely so much on the data which I’m the person to provide the data for and the information.” – HRIO, User Interview 3 Line 10*

The expectations and hopes of this user group are that electronic feedback systems help to hold people accountable and allow them to know on a day-to-day basis what is going on, although most records and data collection tools are still paper based. In that specific context, they mention that even if there was a shift to electronic tools, having the option to print electronic information remains important:

- *“You know if you go to an office of your boss, if something is not written, if you print and sign it, [it] will be received. But electronically, he or she might not be able to see [it].” – HRIO Interview 19 Line 66*

6.3.1.c. Network of Actors, Teams & Champions

The interviews revealed a **network** of individuals and groups involved in designing, implementing, and using the dashboard feedback system. Figure 39 provides an overview of who the implementers and users referred to as being part of the feedback ecosystem to improve QoC.

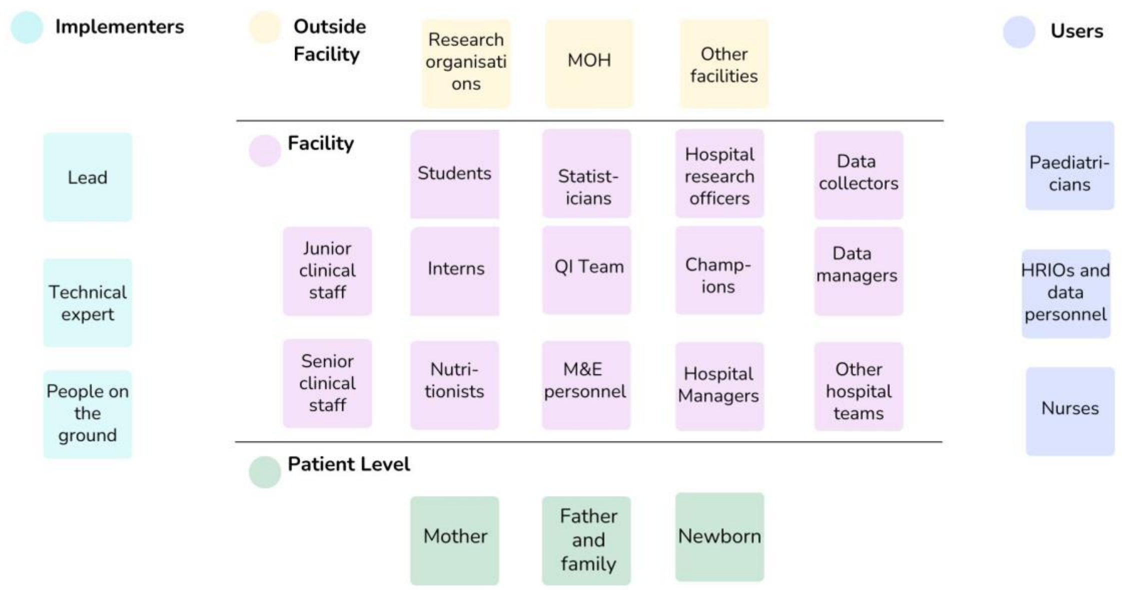


Figure 39: Overview of network actors (middle) mentioned by implementers (left) and users (right).

Some implementers pointed out who they worked with while designing the tool. For CIN this was a rather small team of individuals at the beginning:

> *"Essentially, there was a rather small group of people involved in making the decisions. Central was myself and A* in collaboration with other clinical researchers or clinical team members like B*. Particularly for the neonatal side, B* had a major role." - CIN Implementer Interview 8 Line 18*

In addition to pointing out other individuals and groups, users highlighted the importance of **team meetings** to discuss feedback delivered with the dashboards. When asked about the familiarity with the CIN and NEST dashboards, implementers emphasised that they involved various users in the creation of feedback reports and in discussions on how to improve care quality to drive action.

Implementers specifically mentioned a desire to include leaders from both government and facility levels in routine dashboard processes through, for example, departmental meetings. However, in practice, they explained that it is not always possible to go through the report with everyone, which is why it is important to have **champions** who can synthesise the report for the rest of the team and help share the information either in formal meetings or on a personal basis during day-to-day interactions (Appendix 6.8., Quote 9). Further extending networks, paediatricians expressed a desire to contact teams or individuals from outside their own facility to discuss best practices, saying, *"We wish to be able to talk to other hospitals"* to understand how they, for example, achieved better outcomes (Appendix 6.8., Quote 10). One mentioned calling a medical school friend from another facility. A distinguishing feature of CIN is benchmarking a hospital's performance for a previous period against an average of all other CIN hospitals' performance on an anonymous basis, where participants cannot see which facility is performing better or worse. NEST did not provide any kind of benchmarking at the time of the interviews. Noteworthy is that although users mention the government as a potential user, their impression was that the government was not using data for their decision-making, but rather basing decisions on interests instead of data (Appendix 6.8., Quote 11).

Some implementers shared that they have a **double role** with the CIN tools, acting as implementers and users given their employment with both KEMRI and the hospital as clinicians. One implementer pointed out that they are users of these reports just like the hospital staff, for example, to see what the progress is with an intervention:

- *"In my own thinking and idea, I try to understand that the PI needs to see what the progress [is] and how things [have] been moving before the implementation of [the] intervention Y and after its implementation." – Implementer Interview 10, Line 17*

This confirmed asking, as planned in the interview tool, who respondents think the primary and secondary users are, as described next.

6.3.1.d. Primary and Secondary Users

Each interviewee group, both implementers and users, sometimes referred to each other as the party responsible for using the tools for quality improvements as well as to other more peripheral actors. This raised the question that aligns with the design literature on who the primary and secondary dashboard users are, which is one factor to think about when designing and implementing a tool (see Chapter II).

i. Target Groups the Implementers Had in Mind

Both dashboard implementers report that they had a range of users or target groups in mind when designing the tools, such as the government, stakeholders, hospital management and clinical frontline workers (Appendix 6.8., Quote 12).

CIN implementers considered hospital staff and senior frontline workers (nurses and doctors in hospitals providing neonatal care) as the primary users when developing the tool, despite being aware that the frontline people who see the babies are usually medical interns (Appendix 6.8., Quote 13). The reasons that CIN implementers gave for targeting the paediatricians and senior nurses were pragmatic and about responsibility. Although the interns are the primary caregivers at the wards, they rotate frequently, and so new staff would have to be trained every three months. Instead, the target group was those with the responsibility to supervise and inform the interns based on what they read from the dashboard, as stated by one implementer involved from the beginning (Appendix 6.8., Quote 14). It was emphasised in several interviews that the main user is not an individual but a **team** of senior clinicians who are the primary caregivers at the bedside, which can include other professionals, such as nutritionists. One of the implementers who comes with a double role, being an implementer and user as a clinician, points out that the secondary users at the facilities are hospital administration teams that use the reports to make rational decisions, e.g., about resource allocation (Appendix 6.8., Quote 15).

Besides the hospital management, CIN implementers also had other stakeholders across the health system in mind as feedback target groups, such as the government and researchers (Appendix 6.8., Quote 16). The same implementer added that there are opportunities for many users to build research careers with data available through the network.

NEST implementers similarly had frontline workers in mind. They also had a specific focus on the hospital management team but realised they might not be present most of the time, and so did not consider them as primary users during their implementation process (Appendix 6.8., Quote 17). They have two main user groups: Group 1 comprises themselves as the implementers, including researchers and data people, and Group 2, the users at hospitals that consist of 4 sub-groups: clinicians, managerial and administrative teams, technical biomedical teams and governmental teams (Appendix 6.8., Quote 18). Patients were not considered or mentioned as potential users or contributors by either CIN or NEST implementers. However, a limitation is that I did not plan for this question to be included in the interviews. It was only raised spontaneously in one interview with a NEST person (Appendix 6.8., Quote 19).

ii. User Perceptions Who the Main Users Are

Users were asked about their perceptions of who the main users are. Although not directly asked who the primary or secondary users are, they mention implementers as owners of the tools.

> *"We need training with the use of the tools. You need to provide us with the tools after you've trained us because it is your tools. You can even do [a] pilot study, give us this and see whether they're doing the right thing."*- Nurse User Interview 12 Line 93

This implies the implementers are seen as the primary users, responsible for ensuring everyone can use them, and delivering if necessary. Users also describe the implementers as primarily responsible for setting goals, while the users focus on achieving these goals by improving care outcomes. In these ways the main users are on both the implementer and user sides.

> *Moderator: "Who are the main user groups? You mentioned different actors earlier." Respondent: "Yes, the main users of these reports are both sides, the central CIN side, and that is the PIs, and the other group of users is the hospital teams." - HRIO User Interview 10 Line 9*

On a horizontal level, the users (HRIOs, nurses and paediatricians) do not refer to each other as primary or secondary users, which could stem from their perception that the technical implementers are the primary dashboard users, owning and managing the tools. Meanwhile, they focus on their role-specific responsibilities and tasks around the dashboard and use the feedback it delivers, seeing themselves as accountable for QoC (see Section 6.3.1.c).

6.3.2. Dashboard Design and Implementation

This section presents tool-specific insights from the interviews, including respondents' perceptions, views, preferences, and feelings about the tool's design, implementation, and broader use. These insights help identify successful elements and gaps, informing future tool iterations.

6.3.2.a. Perceptions: What is a dashboard?

All respondents were asked about their perceptions of what a dashboard is in general, its purpose, and whether they knew of any definitions to describe a dashboard. Initially, I included these questions about people's general perception of a dashboard at the start of the interview. However, it was later learned that it was best to ask these at the end to avoid distracting respondents from assessing the CIN and NEST dashboards as specific tools to explore. Users and implementers had similar views on what a dashboard is and what it can offer (see Table 16).

Table 16: Summary of users and implementers perceptions on what a dashboard is

Question	What users think a dashboard is and/or should be like	What implementers think a dashboard is and/or should be like
What is a dashboard?	A dashboard is a **visual tool** that provides a **quick glance of key information**, displays what's happening in terms of information, shows applications and data, summarises activities, and previews data collected in various areas of interest. It is a presentation of key information for an overview, a place to keep data and assess information, and a tool to gather and display information in a summarised way on one page.	A dashboard is a **visual representation of summary data** that can be used to monitor progress, identify gaps, and inform action points. It is a practical tool for quality improvement that can show trends, gaps, or practices. One of the respondents would define it as "*simply a summary of key performance of a project*".
The purpose of a dashboard	The purpose of a dashboard is to **improve the QoC** by providing a snapshot of useful and relevant information.	The main purpose of a dashboard is to **improve data use and quality**, stimulate discussion, generate action, and bring about improvement in a facility.
A definition of a dashboard	Users could not refer to a definition they heard of.	Most of the implementers did not recall any official definition of a dashboard from the literature.

Both groups, users and implementers, could not refer to a specific definition they had come across but provided examples of something they had seen and interpreted as a dashboard. Although there was no specific question about design, respondents shared their thoughts on what a dashboard should be like. Users emphasized that the design of a dashboard should be clear, not overly populated, and self-explanatory. Implementers highlighted that a dashboard should be simple, accessible, and interactive, allowing users to choose what they want to see. None of the users mentioned a specific electronic tool referred to as a dashboard within their facility, aside from CIN and NEST. One respondent mentioned having a dashboard in their facility, referring to a whiteboard displaying trends for the month, and another respondent mentioned a soft copy file with icons showing graphs that summarise data.

6.3.2.b. Implementers' Intentions and Dashboard Development Experiences

When I interviewed the implementers, they described the background, purpose, design, and implementation of the CIN and NEST dashboards, highlighting the feedback mechanisms intended to improve neonatal care. Both tools use data-driven decisions based on KPIs selected from their experiences and guidelines. They also considered the readiness and existing systems to ensure the intervention's integration and effectiveness. Understanding the implementers' intentions and processes helps to identify successful elements and potential gaps, informing future tool iterations.

i. CIN Feedback Tool Development

The CIN feedback tool is used within a network designed to improve healthcare services by ensuring that clinicians adhere to clinical guidelines. When designing the tools, the implementers assumed the tool would be used by hospital staff, particularly senior clinicians who influence junior staff, and that it would improve guideline adherence and, thus, care quality. The feedback was first designed as a written report (MS Word-Pdf document) that presents data analysis results, highlighting specific KPIs to target improvements effectively.

> *"The main goal was to improve the QoC because first of all we started with the indicators where we could look at the level of documentation, [which] was misinformation." – CIN Implementer Interview 9 Line 67*

It uses a colour-coded traffic light system to provide feedback on rated performance. Initially limited by data availability, the CIN network developed a data collection system to feed into the tool. The tool has evolved over the past 10 years to include graphical representations of performance over time, which allows benchmarking with previous performance and other hospitals. As part of their data quality improvement efforts, CIN had to develop their own data system (using REDCap) that was integrated into the Kenyan hospital context.

> *"The first purpose was to… keep track of the services that they're providing and using that information to improve quality, to identify opportunities for improvement, implement interventions and to track how the things are changing over time. The first aim was to work with the hospitals to make the data that they use of better quality in a form that is consistent, and in a form that is useful for them to use in the work that they've already identified as an area where we need data to make decisions. So rather than work with them - since we have the capacity to provide that data - to give them the data in a form that is useful to them. Yes, that's the first aim to support quality improvement at the hospital level in a way that is efficient and makes sure that people don't have to reinvent the wheel …because collecting data is actually quite a time-intensive process."- CIN Implementer Interview 13 Line 10*

When CIN started to develop its network, it tried to integrate its efforts into the existing health information systems but faced challenges quickly. DHIS2 (as the main electronic tool that is used in the Kenyan public sector to track and use data for decision making) did not have individual patient record capability at the time and relied on manual entry of summaries of hospital activities such as admissions. As a result, CIN had to develop its own data collection system based on individual patient data that includes data cleaning and follows standard procedures to ensure data quality.

The implementers acknowledged that feedback had not been taken from the users sufficiently to use the terminology '**co-design**' of their feedback reports / dashboards although that would have been ideal from the beginning. As such future improvements focus on enhancing digital features, narrowing the focus of feedback, and increasing user engagement in the tool's further development. They also acknowledged that a usability test had never been done and they are not sure how well users can interpret data from the existing design, although they collected verbal feedback during six-monthly meetings with hospitals and as part of data quality checks. As such my research was highlighted as a study to address this missed opportunity.

> *“Essentially, there was a rather small group of people involved in making the decisions. Central was myself and in collaboration with other clinical researchers or clinical team members like Y. Our feedback is very much directed at saying whether things in those guidelines are happening. […] We essentially try and give feedback on each of those. And we tried to give feedback on the things that the hospital or the clinicians and nurses should be able to improve.” - CIN Implementer Interview 8 Line 19*

Overall, the CIN tool is integrated within a system of CIN partner hospitals, where data collection and systems have been set up to improve quality data. The implementers attach high priority to data quality and acknowledged that this should not lead to a higher burden of data collection for hospital staff.

ii. NEST Dashboard Development

The multi-country NEST dashboard was created to standardise and improve the quality of neonatal care. The dashboard was conceptualized as an evidence-based tool for quality improvement requiring standardised data accessible at the facility level. For Kenya, NEST data were collected through the CIN data system that was already in place, from a subset of NEST partner hospitals (Appendix 3.1.). The NEST dashboard development involved a systematic process of learning from existing data sets, defining core variables through voting and consensus, and designing the dashboard structure to focus on outcomes, care pathways, and inputs/processes most relevant to NEST objectives. The perception of implementers is that the dashboard enables data use that can lead to actionable insights for improving care quality in a user-friendly 1-page summary format.

The design process was reported to be collaborative (involving multiple stakeholders), influenced by evidence-based practices and adapted to suit needs across multiple countries in which NEST is being implemented. However, the responses about whether it was formally co-designed (or not) were inconsistent, and there seemed to be different perspectives (Appendix 6.9., Quote 20 to 23). Although co-design was seen as essential, the challenge involved balancing the diverse needs of stakeholders, such as clinicians, data personnel, and quality improvement experts. Effective co-design would require setting clear parameters and acknowledging power imbalances, ensuring all groups, particularly the less powerful data personnel, can effectively contribute and manage their roles (Appendix 6.9., Quote 24). According to the NEST implementers, feedback from users is collected to continuously improve and adjust the dashboard design to ensure its usage and effectiveness. From the beginning, quality improvement teams were on the ground conducting health facility assessments to accompany the dashboard implementation process and promote uptake. Throughout this process, the NEST team engaged with the dashboard users, and feedback has been collected to identify areas for improvement. However, a usability test, as incorporated in this study, was not mentioned. Furthermore, although they planned for the tools to integrate with existing health information systems, they face the same DHIS2 related limitations as CIN. They, therefore, "*piggybacked*" on the CIN data system and added quality improvement visits to enhance the NEST intervention specifically (Appendix 6.9., Quote 25). In summary, NEST implementers emphasised the importance of co-design, while noting the challenges of balancing the needs, demands and power dynamics of different user groups. A major ongoing challenge is building data capacity to interpret the dashboard data, moving beyond simply visualising it, and making appropriate decisions based on data quality and plausibility. More details about the challenges will be outlined in the next section.

iii. Observed Tool Preferences

When inquiring whether the implementers could observe users' preferences for CIN or NEST dashboards, implementers were hesitant to conclude but highlighted how their own tool is perceived and used very well. Some reported that users would prefer feedback from NEST because of the prospect of having an electronic dashboard over a written report. At the same time, they clarified that users seem to respond better to spoken information and personal interactions rather than written reports, whether that is CIN or NEST. This suggests that direct communication (person to person) is still necessary even with electronic dashboards, perhaps until the use of dashboards becomes routine for the hospitals. One CIN implementer highlights that users rely heavily on direct communication, which typically occurs through meetings where implementers directly address any identified issues with focal persons at the hospitals:

> *"I think it's the communication. I've realised that they receive information better when it's given to them, spoken information rather than information [that] you have to read. When someone calls you or meets you and gives you feedback on how you're doing, then there's a better chance of them knowing how they're doing that way [rather] than you sending them a report and expecting them to read [it] on their own." – CIN Interview 19 Line 88*

iv. Implementers' Challenges

Besides the challenge of co-designing a dashboard with multiple stakeholders and varied information needs, the implementers faced other challenges, primarily around data quality, the design and implementation of the tools.

While trust in the dashboards by individual users was not the main concern for implementers, they highlighted data accuracy and difficulties in interpreting visuals as significant challenges that can lead to mistrust. Both implementer groups acknowledged challenges related to data accuracy and the complexity of reports and are committed to improving data collection processes to tackle accuracy and reliability, improving presentation to the hospital team in efforts to build user trust (Appendix 6.9., Quote 26).

In addition, there is an effort to design the reports so that they and the specific indicators are easy to understand by providing introductions to the tools and training. Efforts are ongoing to enhance the visual presentation of the tools, including better graphical representations and simplified, colour-coded reports that can be easily interpreted across all prints, but implementers did not report what the strategy for these efforts is apart from having a QI team. These QI teams provide introductions and support in using the tools and also collect feedback from users about the design and content of the dashboards (Appendix 6.9., Quote 27). There is an emphasis on developing more interactive and user-friendly interfaces to facilitate quicker understanding and adoption of insights from the data. One of the implementers who is also a tool user, reflects on how the visualisations were designed and how decisions were made, similarly to how a researcher prepared a research manuscript, meaning the design was created from a research perspective (Appendix 6.9., Quote 28).

How to co-design, deliver training to and engage with different user groups, such as clinicians and administrators, posed challenges due to diverse backgrounds and varying abilities to comprehend the tools. Finding a one-size-fits-all dashboard was reported an impossible task. At the systems level, integrating these tools into existing infrastructures and overcoming technological limitations comes with significant obstacles. As one NEST implementer noted, there is also the challenge of local demand to design and roll out the dashboard quickly. One NEST country pushed for rapid implementation of the NEST dashboard, driven by competition with other departments that already have similar tools. Another challenge is that users want immediate electronic access and linkage to other systems. In Kenya, initial confusion about the NEST's dashboard purpose (the same as CIN's feedback system, to improve the quality of newborn care) led to scepticism at first since it appeared like a project (rather than programme), but a workshop led by the government helped to increase enthusiasm and to see the difference. Nevertheless, the political dynamics of different initiatives and the demand of different actors wanting to go fast with dashboard implementations were seen to potentially distract from the primary goal of implementers to serve facilities effectively (Appendix 6.9., Quote 29).

Several NEST implementers mentioned that there might be plans to integrate the dashboard into DHIS2, driven by the country teams, as implementers are keen to avoid creating parallel systems. Similarly, CIN implementers aim for long-term integration but are sceptical about DHIS2. The CIN collects more detailed information than the national information system. Concerns arise from the interoperability issues and the nature of DHIS2, which is primarily an aggregating system that does not collect data on individual patients and lacks sophistication for inpatient hospital care. One CIN implementer mentioned that while the integration is possible and useful, it is important first to assess the effectiveness of the existing DHIS2 features and how the information is being utilised by its users, which is in doubt it is. Most CIN implementors emphasised the need to improve the CIN feedback system further first, moving towards a more electronic and integrated approach, to prepare for the inevitable transition to EHRs in the healthcare system. They also noted that tackling existing challenges of data flows that already exist with paper-based forms need to be addressed first (Appendix 6.9., Quote 30). Integrating the basic facility-level individual patient data and registers into DHIS2 would be the starting point, as it could allow for better data aggregation and access to information across different departments within the Ministry of Health.

Implementers noted the importance of training programmes to the specific needs of each user group to enhance their engagement and effectiveness in using the tools and how to interpret the feedback, implying that implementers are to some degree aware of user challenges in interpreting the reports. Although implementers mentioned that training was provided as part of the dashboard implementation (introducing how to use it and what visuals mean), there was no "*formal training*" on the competencies such as data literacy and interpretation skills, suggesting there could be training advancements for skills development (Appendix 6.9., Quote 31). Both CIN and NEST provide explanations on how to use the dashboards and interpret content, but implementers emphasised that ongoing support and periodic re-evaluation of the tool's user interface was needed to better align with the varied capabilities and needs across different user groups:

- *"So, I think there is something in training people how to interpret those dashboards in the … of maybe two, three, four times they're given so that people understand it." – CIN/NEST Implementer Interview 6 Line 59*

v. Unintended Consequences Implementers Considered

The potential unintended consequences of the tools were a concern of some implementers. An example is the potential risk that users are discouraged by the feedback they receive, especially during times of staff shortage and high mortality rates (Appendix 6.9., Quote 32). As the goal was that people aim to improve their performance rather than to become demotivated, CIN introduced a RAG traffic light system to their feedback and decided to add a mitigating fourth colour to give "*a little more discretion in the middle zone*", using the colour red for 'very poor' and the colour pink for 'fair', yellow for 'good' and green for 'very good' performance (Appendix 6.9, Quote 33).

CIN implementers further raised the concern of a risk that information from the report is interpreted by the wrong person and a media outlet. This was highlighted as a particular concern for mortality if numbers were used for politics, which might, in turn impact how hospital staff perform and report data:

"If this information is interpreted by the wrong person, especially things like mortality rate, it can be used to castigate the newborn unit team. Yes, so if there's politics involved in the hospital, it can get in the wrong hands, go to the county and the end result is that [they say] 'this hospital

is not doing anything, they're killing the babies'. We've experienced that. Just politics, but not necessarily because it's true, but just because the report was interpreted in a political way to gain political mileage. Again, because there's a lot of politics in the counties. […] Also, when a media outlet wants to report something about maybe our newborn unit, that's the details we provide them with. […] If we don't know what our mortality rate is, that's the data we'll provide them with. So now it depends on how they spin the story, and which politician reads it and decides to use it against us." – CIN Implementer Interview 19 Line 99

NEST implementers mentioned the workload of data collection as a concern in so far as the request for having accurate data requires a lot of work on the ground to generate the dashboard content (Appendix 6.9., Quote 35). For this reason, they considered having additional data collectors, although it seemed unclear how and by whom they are employed. One implementer raised a concern about the consequences of misinterpreting data and of how feedback on performance could potentially lead to "*facility shame"* and, if outcomes are very poor, creating "*a culture of blame",* which is not the intention (Appendix 6.9., Quote 36).

6.3.2.c. Users' Dashboard and Feedback Experiences

All user groups agreed that a dashboard providing feedback on specific newborn KPIs helps identify gaps in care and areas for improvement. The tool helps to assess whether the care provided aligns with expected standards and helps pinpoint specific areas like infection rates and procedure adherence. As described earlier, a narrative talk-through was used with users looking at the visuals during the interview. In the following sections, findings are described, including users' tool preferences, levels of engagement and interactions with the tool, challenges with the dashboard content, and other factors outside of the tool itself.

i. Tool Preferences

When talking to the users, they did not seem to prefer one dashboard tool over the other. Clinical users value both tools for providing a comprehensive performance picture based on newborn KPIs. Some tended to prefer CIN, possibly because they were exposed to it for a longer period or, as they said because it contains more information (several pages). However, the one-page NEST format was also liked, with some voices reflecting that there could be a combination of both tools:

- *"They are both actually important and both have advantages and disadvantages. Do I have to pick one? [...] I would like the CIN one because it's detailed, it has more than just [one page]. For the report its just one page for the NEST and this one [CIN] has more detail. I would not really choose one, but I would say it should be combined. The reports that are more detailed can be put in the CIN one and then the NEST one should just be used to complement the CIN one. And the report on CIN should be the main and then it's complemented on this." - HRIO Interview 3, Line 101*

Users also mentioned that the two tools have their own legitimacy and could both be used for different things, relating that to the frequency of feedback:

- *"They [CIN and NEST dashboard] are talking different things, I cannot say this one is better than this one, so I'll work with the quarterly unital data, and I'll still go back to the quality." - HRIO Interview 6 Line 140*

CIN was appreciated for its encouragement to reflect on past performance and benchmarking with other CIN hospitals, which helps to discuss strategic planning for future improvement with their management, as highlighted by one of the Paediatricians (Appendix 6.8., Quote 10). And finally, the fact that CIN is a report that can be shared seemed to be of particular value, as clinicians use prints or other technologies to share the soft copy of it:

- *"In fact, mostly there is a soft copy. I only send it to the WhatsApp group. The most critical part I always highlight or take a picture and show them how we're performing, so you'll find that people will start improving by themselves." – Nurse Interview 8 Line 67*

ii. Different Levels of Engagement with and Use of Feedback from Dashboards

Each user group has slightly different ways of engaging with and using the tools from CIN and NEST. Paediatricians appeared very proactive and engaged, using the data effectively for strategic decision-making and operational improvements in their roles as team leaders and lined to their managerial tasks. They also seemed to have a deep understanding and were more likely to leverage the data to drive changes within their practices. However, to make better use of the data, they mentioned that the timing plays a role:

- *Respondent: "If they [the dashboard reports] could come in at the same time every month, it would make it easier for us to be able to analyse the data. It also makes it easier for me to do the comparison. If I get the data at the same time, every month probably then analysis, and setting up a time and date to actually go through the data and make sure it coincides with my mortality meeting it will be a bit easier."*
  Moderator: "By the same time, you mean consistency in terms of the 15th every month?"
  Respondent: "Maybe the 15th of the month we get the CIN report so that we know. And if I get the report on the 15th, then by the 18th or 20th, I need to have done my mortality meeting and do comparisons."- Paediatrician Interview 17 Line 112

Most nurses interviewed were senior or deputy in-charge nurses (Appendix 6.4.). They showed variability in their interaction with the tools, often needing additional support to interpret the data, highlighting the need for more straightforward data presentation to use the feedback effectively. Nurses also carry significant responsibility in ensuring other team members understand and act on the dashboards' feedback. Since staff turnover is frequent, they became the trainers of incoming staff, particularly interns and students, which adds to their already high amount of work.

- *"We are sensitising them [on the data]. Today we invited some staff from KNH newborn unit to come and teach the staff. We invited them, and then for us that we know we do it as a* *continuous sensitisation on job training. Especially to our students and interns. You know we have a lot of interns there, so we teach them, so it becomes a continuous process. Yes, sensitising them. – Nurse Interview 11 Line 182*

HRIOs focused more on operational efficiency and resource management, indicating a more administrative than clinical use of the data. This user group was predominantly concerned with data collection, quality and how data could streamline operations rather than direct patient care improvements. The interviewees also touched on the use of CIN and NEST feedback during audit team meetings and Maternal and Perinatal Death Surveillance and Response meetings. All HRIOs mentioned several departmental meetings during which CIN and/or NEST feedback is regularly discussed and data reviewed. Although CIN meetings to discuss feedback are valued, there is a demand for meetings to discuss feedback internally within each hospital:

- *"We need a feedback forum. Within the institution not outside the institution, because when we are seated all of us, we have the Migori [hospital], we have the Machakos [hospital], their problems might be unique to their environment. Ours is also unique to our environment. We need the forum here." – HRIO Interview 16 Line 89*

Finally, the current user feedback system for reporting challenges to implementers relies on users proactively seeking help from implementers, except for when there are QI visits or meetings with the implementers. While users can give feedback during these visits or meetings, they often lack immediate support when using the graphs independently for internal purposes.

- *"Finally, for feedback, feedback that would be two-ways [would be better]. One [reason] is, why agree with the contents of the reports? There is one assumption here, that this report is correct. [But] is there a chance it does not reflect [our performance]? It would help to hear from me what my thoughts are. When I am told that this number of babies died, and I have* *serious evidence that they didn't, is there any way to feedback this and just have a conversation not really to dispute but to have a conversation." – Paediatrician Interview 13 Line 133*

Overall, the interviewees emphasised the importance of collaboration and communication among implementers and the different healthcare professionals and user groups on the ground to improve the QoC in the neonatal units.

iii. Data Quality

All user groups mentioned that the dashboards help to identify gaps in care and areas for improvement against standards, but the quality of data that feeds into dashboards is essential. All users highlighted the interplay between data accuracy and trust and, thus, the effective use of dashboards for quality improvement.

Data Quality, Reliability and Trust

All user groups highlighted issues related to data quality as affecting the way they feel about the feedback (whether it reflects their actual performance) and, thus, their use of the information provided by the tools. Concerns were raised about the source of data collection (patient files vs. registers), as well as the broader challenges with information systems on the ground, such as the source document. Multiple varying data collection tools present different data for the same patient or indicator, which impacts the reliability and trust of the feedback (Appendix 6.9., Quote 37, 38). Mistrust in data is a significant concern, as it influences whether users establish a sense of ownership over the tools and the data provided:

- *"I always capture my information but then to own an information that is written by another person, now it becomes a bit difficult [to trust]." – Nurse Interview 8 Line 52*

Although the complex data system is primarily seen as the reason for accuracy problems, one user noted that mistrust might also stem from the failure of users to interpret visuals correctly and that once the feedback graph is explained, users tend to understand and trust the data more. This understanding then leads to accountability for outcomes and reinforces performance improvement (Appendix 6.9., Quote 39). Regarding data quality, nurses also raised concerns that doctors sometimes make quick, possibly inaccurate diagnoses, questioning the overall data quality:

- *"Also, with that staff there are some people who are very quick. You have seen, like that file I've said that the doctor to go, we have a file, which is like a checklist, but you see, there are some who just go straight to the diagnosis and treatment. Challenges are on both sides. … You know it affects the data because you see now [clinicians] don't do the observation, [because] in the data it will show that with that patient [something specific] was not observed" – Nurse Interview 15 Line 121*

Although users raised concerns about data quality, including for CIN, they still perceived CIN data as more accurate and trustable than from other tools. As a result, the CIN report is used to produce other internal and external hospital reports, with reliance on the data systems created by CIN. HRIOs assume the CIN report is more accurate because their data clerks take data directly from patient files, where checks and balances help identify discrepancies (Appendix 6.9, Quote 40).

Feedback Impact on Data Quality

An important finding was that users are guided by mistrust in data to follow up on discrepancies, which helps improve their facility's data quality by identifying data generation and collection errors. For example, a senior nurse emphasised the importance of discussing data in meetings because *"it might not be easy for you to realise your own data that you're generating is wrong*" (Appendix 6.9., Quote 41). HRIOs also praised CIN for improving data quality and noted that trained officers could apply these skills even when transferred to other hospitals (Appendix 6.9., Quote 42). When discrepancies arise between CIN data and other reports, HRIOs investigate the source of the problem, initiating conversations about improving documentation and recording of data in the hospital (Appendix 6.9., Quote 43).

Data for Advocacy

Users noted that if the data challenge is addressed, it can speak for them and highlight needs. They can use the dashboard to communicate with various departments, advocating for necessary supplies or additional staff. The NEST dashboard is viewed as an opportunity to use data to argue with consultants or administration to find causes of outcomes (Appendix 6.9., Quote 44). However, HRIOs believe that the push for data-driven interventions should come from clinicians rather than records personnel (Appendix 6.9., Quote 45).

Across all interviews, data quality, trust in data, and the importance of data in quality improvement, advocacy and communication were prominent themes. Interviewees felt that data quality is crucial, and that training and interpretation skills could enhance the trust in data and further contribute to data quality, creating a positive feedback loop that supports a comprehensive performance picture and drives continuous improvement.

iv. Visualisations and Interpretation Requirements

The survey results revealed that many users struggle with correctly interpreting visuals, which is why all interviews were carefully screened for explanations. Overall, the challenges of interpreting visuals seem to result from the complexity of data systems, difficulties in understanding complex graphs, and the need for clear explanations to overcome initial scepticism.

Guidance and Training

Users find data analysis challenging and often struggle to interpret graphs without guidance. Once data is explained and contextualised, staff are more likely to accept and use the information as intended. This highlights the need for clear explanations to make data more accessible and useful (Appendix 6.9., Quote 46). Users emphasised the importance of having the implementer who prepared the data to provide an overview:

- *"I want you as the researcher, I want you as the compiler of the report, to work with me. You see the challenges I go through; you offer what we call supportive supervision. So that when you are going to interpret this report, there is going to be no blame game." – Nurse Interview 2b Line 12*

Not everyone can interpret graphs easily, so guided interpretation is necessary, suggesting the currently provided data visualisations are not easy enough and/or training insufficient. It is considered essential to ensure that **everyone** understands the tools correctly. This indicates a need for training sessions to improve data literacy (Appendix 6.9., Quote 47, 48). One senior nurse highlights that the dashboard graph requires the user to perform some calculations (Quote 46), indicating that it requires a senior person who is well-trained in interpreting graphs to train others on the ground.

Simple Graphs

Furthermore, simplifying visuals can make interpretation more manageable, and users find interpreting graphs with multiple overlapping lines time-consuming and frustrating. This points to the need for more straightforward, less cluttered graphs to avoid confusion and enhance usability by avoiding overly complex data visualisations:

- *“The Venn diagram! Even the last time we were in that meeting with the CIN [implementers], it was this one that people talked about. The Venn diagram is hard to interpret.” – HRIO Interview 10 Line 48*

Users may not be well-trained in data analysis, making it challenging to understand all visuals. The challenge to interpret the visuals is “multiplied’ by the complexity of graph designs and restricted time to interpret them, e.g., multiple overlapping lines and data points that are time-consuming to interpret, particularly for users who already have a high workload or who do not have a strong analytical background. Clinicians seem to be fine interpreting simple line charts; however, when the same chart design is used to fit in more information, it becomes too "chaotic" and users “a little frustrated”. Therefore, simpler graphs with less information are preferred over more complex ones (Appendix 6.9., Quote 49):

- *Now, it takes a lot of time to look at this and there are so many [lines in one graph] – Paediatrician Interview 18 Line 121 referring to CIN line chart, see Figure 5.4., Chapter V.*

However, it seems there are varied responses about whether a line chart or a bar chart is easier to interpret:

- *“I think the last meeting I attended they also talked about the bar graph, the line graph, and all that. They gave us time to comment on them and I think people commented and I also commented that bar graphs teach [us] better than line graphs, especially when you are using one line graph to communicate much information. Maybe you have the deaths, you have the admissions, you have all of that in one line graph.” – Nurse Interview 8 Line 141*

Colour-Coding

The use of red indicators as an alert system to highlight dangers for specific metrics is appreciated, as it signals the need for immediate action and priorities to look at:

- *"The colours are okay because you know, when you see the red you see danger." – Nurse Interview 15 Line 102*
- *"I tend to run to whatever is red. I can flip through this report in one minute, but I'll look at the red ones, stop, move, and then now, probably like two or three days later, when I'm seeking to assimilate the information, that's when I go through everything." – Paediatrician Interview 14 Line 47*

While colour-coding supports understanding, there is a need for printable versions that do not rely on colour, as printers often cannot print in colour (Appendix 6.9., Quote 50). Users prefer to print feedback for meetings or management, as noted above. Overall, users initially question the relevance or accuracy of complex data but are more likely to accept it once it is explained. There seems to be a demand for clear guidance, simplified visuals, and standardised training to help users interpret and use the data to improve care under the high workload they already have. Implementers still need to address these needs.

v. Dashboard Content and KPIs

Being exposed to both tools, the CIN and NEST dashboard, users talked about their perceptions and how these influence their intention to act on the information through KPIs delivered with the tool.

Clinical users value the KPIs from both dashboards for providing a comprehensive performance picture. The CIN feedback encourages reflection on past performance. It provides data for strategic planning for future improvements, whether through improving documentation and data quality or advocating for more resources to enhance care for newborns. The insights from the KPIs can then be used to adjust practices and are hoped to improve patient outcomes. Both dashboard contents seem to motivate users by showing the real-world impacts of their work, providing a sense of accomplishment and clear areas for improvement. All users considered the dashboard KPIs as adequate and complete:

- *“You see, with the thing [the dashboard], I think they are capturing everything. I think I have no problem with it.” – Nurse Interview 15 Line 77*

- *“Within ward rounds to my nursing notes to my treatment shifts, to my discharge door. I don’t know if there is anything I would add because it has picked quite a bit. […] The three months report is kind of bulky. But then if I look, I also do know what I can subtract because it is picking much what would be needed. No, I don’t think I would add anything.” – Paediatrician Interview 14 Line 102*

From all KPIs currently presented, users often look at mortality rates first and admissions or staffing second to ensure that outcomes are not attributed to their actions, using the dashboards as a reassurance tool.

- *“So then of course, mortality. That is one page that I really go to because it helps me to understand why these babies died. Like for the longest time, even before I went to do neonatology, that page disturbed me enough to synthesise ‘why am I losing 20 babies in a month and most of them [are] dying within 24 hours of arrival?’ and it led me to conclude that… After the depression of mortality, when I look at how the other CIN hospitals sick babies are doing, [I see] I am not alone. I am not alone, we are grappling with the same issues, the same issues across the country. When I look at how many pre-terms we are losing, and we have the same issues, then so, for me, those are the key highlights of the report that really catch my attention.” - Paediatrician 13 Line 40*

However, if KPIs indicate otherwise, this can be demotivating. The NEST dashboard presents staffing numbers as part of their health systems section (see Figure 25). One user mentioned, confirmation of (low) numbers on nurses being on shift in contrast with (high) mortality rates would provide reassurance that mortality is not due to personnel but a system's failure. In that regard, some users said the dashboards could act not only as a reassurance but also as an advocacy tool for talking to the hospital management about securing more staff and resources, as otherwise, the mortality rates would not be feasible to reduce. However, providing staffing and equipment numbers seems essential to put other KPIs and hospital performance into context and not demotivate users. Furthermore, reporting on clinical documentation as a KPI in feedback was praised by users as an incentive that had already improved clinicians' practices and, thus, data quality through CIN.

Users also made suggestions regarding KPIs that are currently not presented in dashboards that should be included. All nurses highlighted the importance of having more information about the health status of the mothers and how they provided care post-discharge beyond what they currently receive from dashboards. Referrals were mentioned by nurses as an important indicator to be able to follow up on and justify performance when there are many referrals of very sick babies:

- *"We find most of these referrals, most of them reach to us when they're very sick. This data helps us. When you know this baby was referred from which facility, we can also be able to identify those facilities that referred to us, and at least you try and reach them [to] discuss some issues." – Nurse Interview 11 Line 44*

And some users also mentioned that having feedback on the whole departmental performance might be useful to work on improving care.

vi. Feedback Delivery Timing

As the tools are seen to add value to discussing performance during meetings, there could be adjustments in terms of the timing of delivery of the feedback with both or either of the dashboards (see, Section 6.3.2.c). Participants mentioned various meetings, such as monthly management meetings, bi-weekly audit meetings, and frequent ad-hoc meetings, that were incentivised and supported by the feedback from dashboards.

Monthly management meetings, for example, require detailed reports to support managerial and clinical decision-making, addressing topics like referrals and staffing needs. The bi-weekly meetings often include mortality audits for more frequent reviews and action plans. Receiving frequent feedback was considered to help guide immediate action by users, with red indicators, for example, highlighting urgent issues that require immediate attention. If this feedback is delayed or retrospect, users say it is almost impossible to change outcomes. While quarterly and monthly feedback is appreciated, there is value in real-time or at least weekly data updates:

- *“Because again, data is very important in requesting change. There is that data that you can get in one week, it prompts you that if we continue this trend, we might be in the bad side by the end of the month. There are those easily adjustable challenges that you can do within a shorter time instead of waiting for us or still waiting for maybe CIN to come and tell you this is how you've performed in the last month.” – Nurse Interview 1 Line 90*

Users also suggested different dashboards for different needs: detailed reports (like CIN) for monthly and quarterly meetings and shorter, alert-focused dashboards (like NEST) for immediate attention.

6.3.3. System Integration and Contextual Influences

This section presents the influences specific to the wider health information system and other contextual that respondents mention as impacting the tool and data use, and efforts to improve the QoC.

6.3.3.a. Integration into Existing Information Systems and Practices

i. Information Systems

Existing data systems, including DHIS2, are technically complex and often not interoperable. The complexity of data collection and systems in place has already been mentioned as part of the reasons for trust difficulties based on data quality by users and the difficulties implementers faced when they tried to integrate feedback systems into wider information systems. DHIS2 was mentioned by both implementers and users as the information tool used by the county and national government, and thus, potentially the tool feedback systems could be integrated in the future.

However, a specific issue users highlighted is the variability in source documents. In addition to the technical challenges, the aggregate-level systems in place must handle data from multiple sources and navigate the structural challenges. These source documents often lack consistency, leading to difficulties in data accuracy and trust. For instance, when new staff members, such as paediatricians, join, they may not know the appropriate documents to use, highlighting the need for better training and standardised procedures:

〉 *"The source document that the clerk use, I think that is the problem. In the file it is written, that one I can vote for it, because it's during my ward round. But where this clerk gets the information is what I need. Although I have seen here, we have…for the hospital it is just a general document. […] Even if a new paediatrician comes, they will not know if that is where you need this information to be picked from. Where the clerk picks the information from is key and* *important to us. It is something that also needs to be enhanced." - Paediatrician Interview 17 Line 27-30*

Furthermore, users recognised the potential of integrated feedback systems that link different hospitals, allowing for comparative analysis and better resource allocation. However, there is awareness that implementing such a system involves significant IT-challenges:

- *"I think if there was a certain database, in that I don't know if that can be done, but if we can, [it could] be linked to other hospitals which are doing the same thing. Although that could be an IT issues all together and we are not all tech savvy. But if let's say we can be getting a report from different hospitals and just have a summary of ours, we already have that, I think like when you compare the two another hospital and how they are doing compared to you it's there." - Paediatrician 18 Line 33- 34*

However, this integration could facilitate not only the generation of summary reports from various hospitals but could also enable clinicians and managers to understand their health system more broadly and potentially inform the allocation of resources, for example to a referral hospital.

- *"Maybe we submit our report we say, 'we are having so many referrals from this specific facility and these referrals come because of a specific problem'. So, interest comes, what is it that this facility is not doing they end up referring, referring, referring?" – Nurse Interview 1 Line 38*

However, these data and information systems complexities underscore the need for better data management strategies and robust IT support to optimise the wider health information systems.

ii. Managerial Structures and Clinical Practices

Users highlighted the complexity of clinical practices and hospital management when talking about the use of data to enhance newborn care quality and prioritise key indicators. They noted that the discussions on performance improvement are best integrated into departmental meetings, where dashboard KPIs can be discussed and inform planning.

Vice versa, practical involvement in improving the dashboards was suggested, such as having clinical hospital representatives attend QI meetings and encouraging implementers' participation on the ground. Thus, new tools, such as the CIN and NEST dashboard, need to integrate into existing structures of clinical governance and management.

6.3.3.b. Resource constraints and organisational support

The effectiveness of feedback tools was notably impacted by organisational support and available resources and procurement. Nurses and HRIOs indicated that lack of adequate or inconsistent resources were barriers to improving performance (Appendix 6.10., Quote 51). Although the dashboards were mentioned as advocacy tools by giving them the evidence and having legitimate requests for more support, whether this is for more staff or equipment, there are still many constraints that make it hard to improve performance. Users highlighted the need for more organisational support, noting that feedback often reveals outcomes that can only really be improved by system managers and providing additional resources:

〉 *"The congestion is there, the sharing of couches and incubators, to act for babies maybe in an incubator, and then using a shared bed because if the numbers go very high, we use the [inaudible?] bed. The infrastructure is not good." - Nurse Interview 8 Line 105*

When asked about other missing resources, users mentioned shortages of staff and stationery. Missing supplies, maintenance and stationery add to the concern of documentation and data quality on the patient management nurses and doctors provide (Appendix 6.10., Quote 52 to 54). As mentioned in Section 6.3.2.c, printer issues (black-white prints) can also lead to misinterpretation of colour-coded data, while printing dashboard feedback and paper formats seem to remain important for users to use feedback for advocacy and during meetings; for example, when discussing hospital referrals (Appendix 6.10., Quote 55).

6.3.3.c. Staff turnover and shortage

For users, the main contextual challenge to drive change based on the feedback is the quick turnover of staff, high workload and staff shortage. Nurses particularly reported the burden of workload. This issue not only affects the ability to act on dashboard insights but also impacts data quality and trust in feedback. One nurse noted, "*these reports they will speak for us to advocate for more staff, but remember we are living in a Kenyan world where staffs will never be enough*" (see Quote Nurse Interview 2b), indicating that unresolved workload and staffing issues can potentially demotivate staff and reduce any impact of the dashboards.

Nurses highlight the difficulty of implementing changes due to high staff turnover, which means that by the time monthly reports are received, the team responsible for interpreting and acting on them has often changed. This turnover creates a continuity problem, making it hard to sustain improvements. The frequent rotation of medical interns, who are often unfamiliar with previous data forms and collection, complicates the process of improving data quality and driving behaviour change based on the dashboard feedback (Appendix 6.10., Quote 56).

Nurses further reported on the extreme workload conditions as their general main challenge, where a handful of nurses are responsible for a high number of newborns, leaving no time for breaks, data discussions or recording (Appendix 6.10., Quote 57, 58). This reality forces staff to adapt constantly, often compromising on essential tasks due to overwhelming workload, and the timing to record data and complete information requirements. The struggle with documentation due to high workload and staff turnover forces them to prioritise only the most critical records (Appendix 6.10., Quote 59, 60). The role of the dashboards as an advocacy tool is acknowledged in this regard. Data can highlight needs and communicate to various departments, advocating for the provision of necessary supplies or staff (Appendix 6.10., Quote 61). And despite all challenges and high workload, clinicians are very appreciative of the feedback and willing to make time to improve care:

- *"CIN has revolutionised newborn care. Not just newborn, starting with paediatric, I think country wide. At least it's no longer a secret that paediatrics... It's because we have data, we can churn out data we can speak authoritatively, and a lot of things, this is thanks to CIN. Thanks, that's why I will always give you time." – Paediatrician Interview 13 Line 147*

In summary, these insights paint a picture of a challenging environment where staff are stretched, turnover is high, and the ability to effectively collect and utilise data is affected by these challenges.

6.4. Chapter Summary

6.4.1. Strengths and Limitations

Although using both quantitative and qualitative data added significant strength to this study, qualitative research also presents its own challenges, which shall be reflected here.

Considering social desirability bias when conducting interviews is important to reduce the risk of participants responding in ways they think are socially acceptable and expected, or simply wanting to be polite by saying "*everything works well*" [212]. I was aware of the risk that I am seen as a dashboard funder or linked to the collaborators CIN and NEST in some way, which might influence responses to my questions. Also, while the RA was Kenyan, she was also a CIN employee. We explicitly highlighted the study's purpose to participants and our roles as researchers with an objective stance, trying to understand their needs as users of these tools and the data to improve them. In general, I was considered an outsider, to the country and dashboard programmes, which seemed to encourage participants to discuss their challenges openly. However, I checked the interview tool for neutral, non-leading language and debriefed with the RA after 2-3 interviews to ensure the questions were clear and minimise bias.

Purposeful sampling is an appropriate sampling approach for qualitative research as it supports gathering information-rich and relevant insights from populations and sub-groups of interest [126]. Clear criteria and reflexive practices were employed to approach diverse participant profiles (e.g., variety of professions and hospitals), and to explore the full range of perspectives from relevant stakeholders [213], [214]. The exception was for people from the Ministry of Health. Although there had been attempts to interview government officials, interviews were not feasible due to time constraints and the high workload of potential participants during and after the COVID-19 pandemic.

However, with 42 content-rich interviews of key target participant groups and having reached a point of redundancy with my questions, my sample size was appropriate to explore hospital dashboard usage in Kenya. The answers to the ‘talk-through' on how tools are used in daily practice evolved into asking people what they look at first when they receive information through the CIN or NEST dashboard. As a result, the responses gathered during the study could not be used as an equivalent to the user journey mapping exercise as initially planned. However, it provided valuable information about the graphs and user interface for the CIN and NEST dashboards.

The similarity of findings across users and implementers may be due to the interview process and questions through the interview tools steering responses towards selected themes or viewpoints. To reduce this risk, both interview tools were designed in line with the literature and reviewed by senior researchers. Initial (pilot) interview transcripts were carefully reviewed to reflect and adapt the interview tools and approach (e.g., encourage participants to talk freely) before continuing with other interviews.

Due to COVID-19 pandemic restrictions, virtual interviews may not have captured the same depth as in-person interviews with implementers, however, it provided greater flexibility and an opportunity to talk to individuals who may have otherwise been unavailable. Despite the risks that are common in qualitative research and that were minimised through a thorough approach, the qualitative component provided a strong body of qualitative data with rich findings that complement the quantitative findings presented in the previous chapter.

6.4.2. Dashboard Use: What People Say, Think and Feel

The qualitative exploration highlighted the challenges and considerations from the perspective of users and implementers. These findings added valuable information and the real-world challenges that the quantitative survey with a focus testing interpretability only would have never discovered.

Users reported a range of thoughts, feelings, and experiences, from clear preferences to indifference regarding the tools used to deliver feedback on performance. The analysis revealed several results regarding users' interactions with the dashboards, adding depth to the survey results about what people can (or cannot) do by reflecting on attitudinal, behavioural and contextual factors. The findings of what users say, think and feel are mapped in Figure 40. and provide foundational information for guiding a re-design.

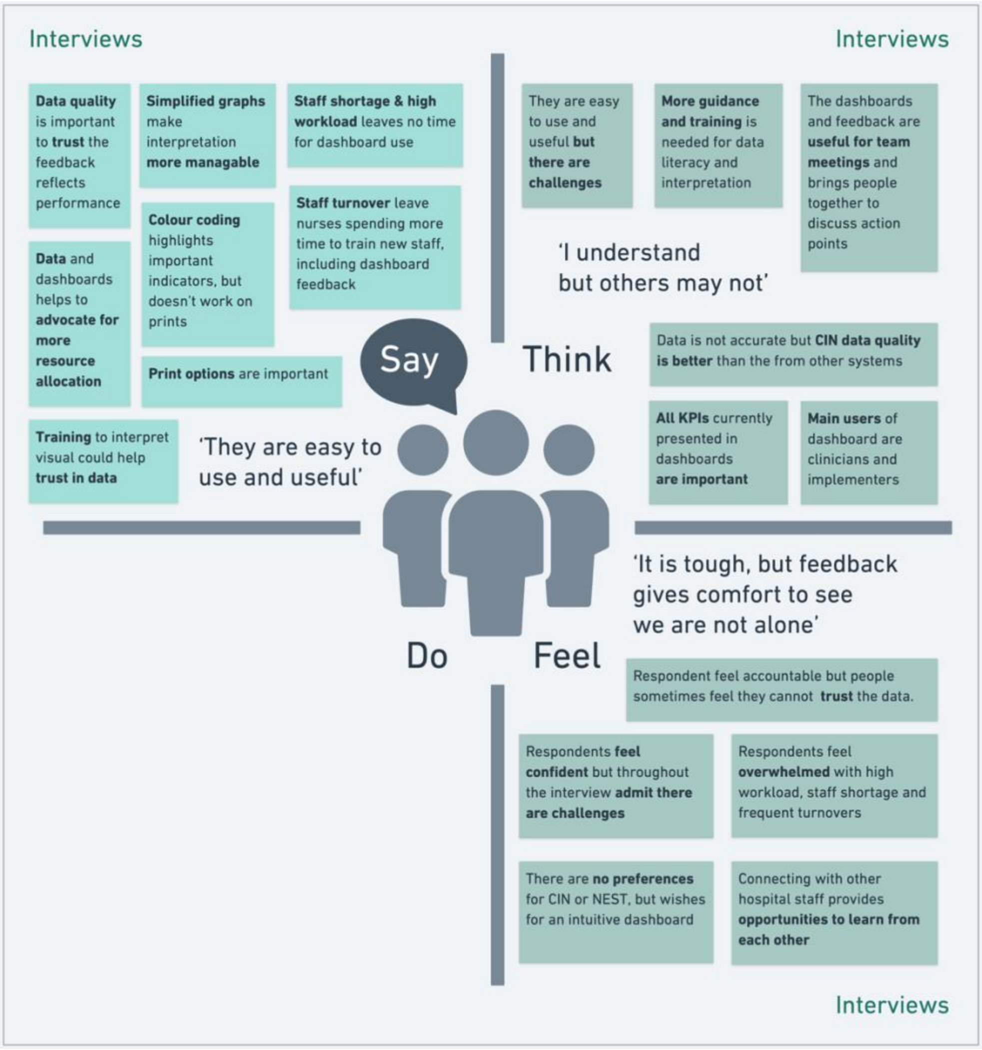


Figure 40: User empathy mapping exercise using qualitative data – What people say, think, feel. Created by CH, adapted and inspired by the Empathy Mapping by NN Group [89].

In summary, the respondents' perceptions indicate that while the CIN and NEST dashboards are valuable tools supporting clinical and managerial tasks, challenges linked to trust in data and the complexity of health information systems, clear data and information communication, staff shortages, and training needs remain critical. Further, the findings underscored the importance of considering active user engagement, straightforward data presentation, and organisational resource support when implementing data-driven feedback tools in healthcare settings. These factors critically influence how dashboards can drive improvements and achieve intended outcomes.

Implementer interviews reveal several insights into the design and implementation intentions and challenges behind the feedback tools. Despite differences in design—such as a one-page dashboard (NEST) versus a multiple-page, report-based tool (CIN), and electronic versus hybrid formats, there are more similarities than differences in the implementers' intentions and goal setting for these feedback tools. Both dashboards were planned to be continuously refined based on user feedback. However, it was not very clear how this process would work to ensure that the tools remain relevant and effective in addressing the evolving user needs in the future. In this respect, there were also varied responses and perceptions regarding whether the tools were directly co-designed (or not). Both implementers regularly incorporate verbal user feedback (from meetings or hospital visits) to improve the dashboard, thus, they consistently engage with users. However, a two-sided feedback system integrated into the dashboards was not considered. Overall, both implementers aim to support QoC improvements for newborns in Kenya through their tools. Developing a strategy for better dashboard designs could help them achieve this goal by shaping effective and impactful digital solutions for care improvements. Table 17 provides a summary of user and implementer findings based on the three domains people, tool and system, categorising different UX and interaction factors identified from the qualitative data. In the next Chapter VII, I will bring the findings from the qualitative and the quantitative data together. After the integration I will then discuss both findings regarding the literature.

Table 17: Summary of findings from user and implementer interviews

Domain	Dashboard interaction factors	User Interviews	Implementer Interviews
PEOPLE	**Skills** and training are required for interpreting data visuals	Users have mixed feelings about their ability to use dashboards and being able to interpret visualisations. Many expressed a need for more training and guidance, particularly in data literacy. Some pointed to others not being able to interpret.	Implementers acknowledge the importance of providing explanations and visuals to enhance users' understanding and use of dashboards. Both implementers say they already provided training; however, they never tested interpretation and usability, or users' data literacy.
	Knowledge exchange helps to learn from feedback	The ability to translate information into action varies with users' roles. HRIOs believe clinicians should lead quality care improvements, while clinicians feel they understand the data better and must push others for data quality. Team meetings (also across facilities) are considered important for knowledge exchange.	Implementers understand the importance of different user roles and knowledge and support them with additional meetings (which provide exchange opportunities with other hospital colleagues) and personnel (e.g., data clerks).
TOOL(S)	**Information** visualisations and dashboard formats affect how	Users appreciate that both tools cover essential KPIs. For daily use, they prefer straightforward graphs containing less information and a one-page format. For monthly meetings, they favour comprehensive information tools such as the CIN report. Users look to implementers for explanations and effective visuals, indicating a reliance on	Implementers work on integrating user feedback regularly, and tools have been updated a few times based on feedback. But they never tested which graphs work best or better, and designs are based on research practices, however, not on best practices from the design and information science literature that provide guidance on simplicity.

	people use data	external support to enhance their understanding and use of the tools.	
	Attitude positively affected by the tool	Overall, users have positive attitudes towards the dashboards and appreciate a reflection on their team performance. Benchmarking and trends can provide a sense of accomplishment. Numbers on staffing and equipment reassures users in times of high mortality rates that it is not individual care failure.	Implementers carefully considered potential unintended consequences and how to provide a tool that supports QoC efforts. Initially, they worried about negative reactions to high mortality rates but found that users remain motivated by the data to drive improvements.
SYSTEM(S)	**Trust** affected by complex systems and data quality	While the CIN system is trusted more than others, concerns about underlying data quality persist. Nurses and clinicians are particularly worried about data collection accuracy as it also directly affects clinical practices and outcomes. At the same time, users trust data enough to use dashboards for advocacy and system changes.	Implementers seem reluctant to believe that there could be trust issues linked to the dashboards since users say they are useful. However, they shared concerns linked to data and information systems and addressed these by working on better underlying data systems. They also try to integrate into existing systems but find it challenging.
	Behaviour affected by systems factors	Users often focus on mortality data but struggle to interpret process indicators, which are important for immediate action and behaviour change. They consider team meetings to discuss next steps as a crucial part for driving action with support by the dashboards feedback. Team dynamics and broader system elements (e.g., workload, turnover rates and other resource constraints) heavily influence how users can act and use dashboards.	Implementers know the resource constraints in the general LMIC hospital setting and how they affect performance. The dashboards are only one part of providing additional support through research, technologies and equipment. They acknowledged the workload burden and support with additional personnel, particularly for the tool implementation and maintenance. They also support user engagement in team discussions by setting up feedback meetings.

Chapter VII. Synthesis and Discussion

This chapter presents the mixed methods integration and discussion of findings. After a brief recap, Section 2 integrates the quantitative and qualitative components, and Section 3 discusses the empirical findings alongside the scoping review. Section 4 reflects on the study's findings in relation to the broader literature, including design principles, and outlines the study's strengths and weaknesses. Finally, I provide recommendations for the design and implementation of dashboards, and future research.

7.1. Introduction

This study aimed to answer how dashboards help to use data to improve the quality of neonatal care in LMICs, with a specific focus on the Kenyan health system. A central aspect of this is understanding how dashboards need to be designed and implemented to contribute to data-driven action. As explained in more detail in Chapter 3, a mixed methods approach combined with extensive background literature provides a comprehensive foundation to assess the complex interactions between people, dashboards, and systems in the context of improving the quality of newborn care in Kenya.[215].

Brown et al. suggest that 42 high-confidence hypotheses, based on 65 qualitative A&F studies, can be viewed as recommendations for feedback to be effective. The authors claim they can serve as guidance when testing feedback interventions and help explain why some interventions may not be adequate by identifying weak points, barriers, and potential facilitators [55]. Neither CIN nor NEST implementers mentioned using CP-FIT specifically as a guidance tool. However, many elements that address CP-FIT hypotheses could be found (Checklist, Appendix 7.1.). The survey showed that users report both dashboards as easy to use and useful, and yet about 50% of the respondents failed to interpret at least half of the neonatal KPIs they were tested on from both dashboards. With the survey complemented by interviews and vice versa, I could explore contradictions that otherwise would have remained hidden.

7.2. Mixed Methods Synthesis: Integration & Interpretation

This section integrates the quantitative and qualitative findings (see Table 18 below), interpreting them alongside Brown et al.'s 42 CP-FIT hypotheses [55] (Section 7.2.1. – 7.2.3.) and expanding the theory with the broader design literature (Section 7.2.4.) to provide a comprehensive view of the factors driving data and dashboard use to improve newborn care quality of newborn care.

Table 18: Recap of f*indings from each data set* using different methods and participants

SURVEY	**USER INTERVIEWS**	**IMPLEMENTER INTERVIEWS**
Captured participants' demographics and perceptions, providing a diverse user background. Assessed participants' ability to interpret dashboard graphs, revealing a discrepancy between their positive attitudes towards the tools and their ability to interpret the content, as well as different user groups having a different skill set to interpret them and training needs that influence the use of these tools.	Confirmed most of the positive attitudes. Users value the tools and consider them as theoretically useful. Added explanations, user needs and possible factors for usability difficulties: system integration, complex data and information systems, time and resource constraints, staffing challenges and workload. Users provided feedback on interpersonal and contextual factors impacting the dashboards' ability to help improve newborn care.	Provided additional insights into the challenges of being the architects of dashboards that are genuinely meant to serve users. Added information on the efforts to develop useful tools to improve care. Implementers considered many 'best practices', even if not guided by all of them, and were aware of the associated contextual difficulties. However, they acknowledged they never tested dashboards and assumed users understand them because they said so.
⇒ Identified issues regarding graph interpretability, user interface and overall usability.	⇒ Added contextual factors, individual and group user needs and overall user experience.	⇒ Added technical knowledge of implementers that shape A&F dashboards.

7.1.1. People: Mapping Interpersonal Dynamics and Roles

Here, I reflect on the people-specific findings (see Section 6.3.1.) alongside the 4 CP-FIT **recipient variable** hypotheses 20 - 24 (Appendix 7.1.). People often know what needs to be done, and the success of interventions can be heavily influenced by a combination of individual characteristics and group compositions, roles, and by their attitudes and behaviours. Understanding these as factors of the success of an intervention, with insights from behaviour change theories, is considered important for effective and targeted intervention strategies [55], [216], [217]. Mapping out the roles and dynamics of the people who use the CIN and NEST dashboards helped to identify some of these factors.

7.1.1.a. Health Professional Characteristics

According to the CP-FIT hypotheses 20 to 22, feedback is most effective when directed at health professionals with positive beliefs and attitudes about feedback (20), strong capabilities in quality improvement (21), and significant expertise in the relevant clinical area (22) [2], [5], [6].

The quantitative component provided a broad overview of participants' demographics, their exposure to the dashboards, the use and preferences of tools, and their confidence levels. The high confidence people have in their ability to use dashboard data effectively and improve QoC also came across during the interviews, except for one user who worries that feedback could be misused in politics and by the media as an unintended consequence of the tools. CIN and NEST users predominantly showed positive attitudes in the survey and interviews towards both dashboards and towards the implementers. The survey showed a strong tendency of users preferring digital platforms over the currently used paper-based reports (see Section 5.3.2.a). Yet, in interviews, they shared their wishes for printed report options.

CIN implementers targeted senior nurses and paediatricians in newborn care for feedback, as they can distribute the information best to the rest of the team. In the survey the clinical user group showed better graph interpretation skills compared to the data and hospital manager group. However, they expressed concerns about their high workload, limited time to interact with the tools and limited time to explain the dashboard content and share feedback. Despite their confidence and valuing both dashboards, I found no correlation between confidence and performance in interpretation of graphs from the quantitative data, however, a negative correlation between expressed training needs and interpretation success, suggesting some users may be overconfident. The qualitative data confirmed a desire for more training and support in graph interpretation and dashboard use, highlighting the implementers' responsibility to address these needs.

Providing human support to utilise the CIN and NEST tools (champions and training) as well as team discussions (as a hospital team and beyond) on how to use the feedback information, including the visuals, to improve performance, was something that came up in almost every interview as important for users. This finding was combined with users identifying the implementers as the party accountable for such training and support, suggesting that users wish the dashboard use to be more tailored to their needs and capabilities as users. The request is basically to provide training for data literacy skills and a ‘dashboard customer support service’ that includes a way for the users to give feedback too.

7.1.1.b. Behavioural Response

The CP-FIT hypothesis 23 states that feedback recipients are responsible for two main actions: correcting and improving patient care, and organisation-level changes like medication storage. Organisation-level actions, according to CP-FIT, often lead to greater clinical performance improvement by augmenting the environment and thus enabling multiple patient-level care behaviours [55], [217]. Interviews revealed that users, after receiving feedback from

CIN or NEST, not only implement patient-level improvements but also engage in efforts at the organisational and hospital level. For example, they improve documentation and data quality, advocate for workforce development, and set up meetings with other hospital teams. These actions demonstrate the potential for feedback to drive system-wide changes in addition to individual patient care improvements. Both dashboards can also positively influence team dynamics, particularly by nudging discussions on quality improvement. For example, it helped to raise awareness of the data quality that led healthcare staff to follow up and promote more accurate data practices.

7.1.2. Tools: Understanding Interactions between People and Dashboards

This section focuses on the tool-specific findings (see Section 6.3.2.) and the interactions between people and the dashboards, examining how the tools and feedback are perceived, understood, and used alongside the 19 CP-FIT **feedback variable** hypotheses [55] (Appendix 7.1.). This includes reflecting on the causes for user errors when interpreting the dashboard graphs.

7.1.2.a. Feedback Goal

According to the CP-FIT hypotheses 1 to 3, feedback goals need to be important, controllable and relevant to feedback recipients to be effective. Controllability means that goals are perceived to be within users' control, which is important for users' acceptance and intentions, and thus for feedback to be actionable [55], [218].

The survey and interviews showed that users perceive all KPIs currently presented by the dashboards as relevant and essential to know for improving their performance. However, all individuals and groups involved in the interviews acknowledged that their degree of control to improve performance and QoC is challenging. Although users are confident in knowing how they could clinically improve patient outcomes, they mentioned that these goals are hard or impossible to achieve due to uncontrollable contextual factors such as staff shortages and resource constraints.

While the implementers and users agree on the overall goals, there is more need to explore which ones are controllable and to acknowledge the external and internal factors that are uncontrollable from a user perspective. This could involve aligning goals with users' roles and perceptions of what is achievable, as well as considering the contextual challenges and how these can be mitigated.

7.1.2.b. Data Collection and Analysis

According to CP-FIT hypotheses 4 to 7, feedback is more effective when data collection and analysis are automated rather than manually conducted by recipients, as this reduces the burden on them and increases the likelihood of feedback acceptance [219], [220]. Trust in the accuracy of these processes and the ability of recipients to exclude irrelevant patient data are also important factors for ensuring that feedback is trusted, relevant, and actionable [221], [222].

Accuracy plays an integral part in the CIN and NEST dashboard use and is one of the most robust findings from this study. With the term ‘accuracy’, CP-FIT refers to the data collection and analysis processes that produce feedback and truly represent clinical performance [55]. Both implementer and user interviews identified hospital ‘data quality’ as a significant challenge, which is why CIN implemented their own data collection and analysis system which NEST uses as data source too. However, data quality goes beyond accuracy and consists of 6 dimensions: accuracy, completeness, consistency, uniqueness, timeliness, and validity [223], [224]. CP-FIT addressed two of these dimensions: Accuracy and timeliness as part of the credibility and actionability mechanism. CIN developed and still maintains a data system that goes beyond accuracy, addressing the challenge of improving data quality in all dimensions. But despite the CIN data system efforts and users valuing what it has already achieved, issues persist largely due to systemic health information challenges, which still rely heavily on a complex system and paper-based forms [15]. Users view data from the CIN database – which is collected from inpatient medical records - as of high quality and generally trustworthy, and yet, perceive data as inaccurate or not a true representation of events at times. Users manage these concerns by following up on data sources or discussing issues informally with peers and implementers; thus, the mistrust in data drives improvements in the data and recordings themselves, however, it can risk distraction from clinical tasks at hand.

The **exclusion** of irrelevant data from dashboards and **data automation** was not directly mentioned and was not a feature at the time of this study [55]. Users currently cannot exclude data, but implementers exclude outliers as part of their data quality checks. Progressing toward a fully digital dashboard, it is expected to potentially enhance data automation processes further if best practices and iterative designs are used.

7.1.2.c. Feedback Display

CP-FIT's feedback display hypotheses 8 to 15 state that feedback is more effective when it highlights areas for improvement in recipients' performance level, uses patient lists for transparency, focuses on individual responsibility (specificity), is timely, includes comparisons (e.g., past performance, benchmarking), prioritises key content and employs user-friendly designs (usability), while evidence on the impact of framing and targets was minimal [55].

Usability, as in a user-friendly feedback tool design, ensures that users can easily navigate and interpret the feedback (hypothesis 15) [55]. Implementers said that the tools were designed with the user in mind and with an effort to make them as user-friendly as possible. Despite high Likert scale ratings for ease of use and usefulness, the survey showed that at least 50% of users struggle with interpreting the content; thus, they cannot use the data from tools effectively. The interviews revealed that users refer to themselves as knowing well how to use the tools or improve care but pointed at others for not having the same capabilities. Alternatively, they pointed towards external factors that would influence their abilities, e.g., insufficient support or resources. However, at the time of the survey testing, all participants were at a conference outside their usual hospital environment and were given more than sufficient time to complete the survey that tested the interpretability. At a later stage of the interviews, and walking through the tools, interviewees became more comfortable in confirming that they sometimes cannot understand what is presented. Users also said they could observe how when their peers cannot interpret the graphs, they tend to refer to the data as possibly incorrect. Thus, the mistrust people have in data may be linked to people not being able to interpret and use the dashboard effectively.

Performance level indicators communicate areas where the recipients' current performance can be improved, and **specificity** refers to feedback that targets the performance of individual health professionals rather than their broader team or organisation (hypothesis 8 and 10). **Patient lists** provide detailed information on the patients used to assess clinical performance, meaning they detail the patients included in the clinical performance calculation, which facilitates verification and helps healthcare professionals to understand sub-optimal care of specific cases, take corrective action and learn for the future (hypothesis 9) [55]. In both dashboards, individual patients are not identifiable. One reason implementers focus on the organisational level is that patient-specific data could easily create a negative workplace culture. NEST implementers for example, highlight their efforts to avoid creating a "*blame game*" culture. Users highlight that they know CIN collected data directly from patient files, which is why they trust the CIN data system more than others. However, future options might be to create individual feedback with sole access by the individual.

Timeliness, meaning recent data is used to calculate current performance (hypothesis 11), is something interviews suggest could be improved in both dashboards. Feedback reflecting the most up-to-date information enables timely interventions. Both dashboards present retrospective rather than real-time data, making it difficult from a user perspective to act on the feedback promptly. However, the survey data collecting users' preferences on feedback frequency showed a preference for more rapid feedback for individuals, but organisational and team feedback seems to be acceptable on a monthly or quarterly basis (see Figure 32).

Trends, benchmarking, and **prioritisation** are highlighted as important for feedback to drive action (hypothesis 12, 13, and 14). **Trend** refers to comparing current to past performance, helping recipients see progress over time. **Benchmarking** compares the recipients' performance with standards or others (e.g., individuals, teams, hospitals), to motivate improvement through comparison. **Prioritisation** highlights the most critical areas, helping recipients focus on key improvements to do [55], [225], [226], [227].

The NEST dashboard incorporated monthly trends for admissions and mortality, and CIN included trends for these two KPIs and several others, such as for example for disease patterns, mortality by birthweight, and antibiotic dosage errors. CIN furthermore designed a benchmarking traffic light system that allows users to compare their current performance with their past performance and with all other partner hospitals (as a pooled mean). User interview findings indicated this to be beneficial, and clinical users mentioned that they would like to know more about their peer-hospital performance as it helps them to identify how to improve care from positive examples of hospitals performing well. Users prioritise feedback by quickly "*flipping*" through the feedback to identify the most critical information, often starting with mortality rates. NEST picked up on that and presents mortality and admission as the top part of the dashboard. CIN goes further by highlighting those areas that require attention with the colour-coded traffic light system. Users confirmed the colours to be very useful as they alert them about the urgent areas requiring their attention. However, the value of colours can be undermined if hospitals are limited to black-white printing options and users rely on printouts to make use of feedback.

7.1.2.d. Feedback Delivery

CP-FIT hypotheses 16 to 19 suggest that feedback is more effective when recipients perceive it as supporting positive change rather than punishment, when a knowledgeable and skilled source delivers it, actively reaches recipients (pushed, not pulled information), and is delivered to a group of people. From the interviews, the dashboard delivery mechanisms for CIN and NEST are perceived as primarily supporting rather than punishing the users; the expertise and effort of implementers were appreciated and the wider programmes valued. While both dashboard functions are seen as supportive rather than punishing poor performance, confirming acceptance and compatibility, the word "*blame game*" came up in interviews by both users and implementers as something everyone seems to be aware of as a potential risk of feedback. The "*push*" delivery method is appreciated, although the survey and user interviews indicate there could be more timely feedback. In interviews, users confirmed that face-to-face interactions from time to time are highly valued as a delivery method and that printout options of the feedback are also helpful, highlighting the need for flexibility to adapt delivery mechanisms when moving towards digital feedback systems. The dashboards are also perceived as delivered to relevant groups within the hospital and as a driver for teamwork and collaboration from the user perspective.

7.1.3. Systems: Exploring both Dashboards in Specific Context

Here, I focus on how the CIN and NEST dashboards fit into existing systems and context. This links to the system-specific findings (see Section 6.3.3.) alongside the 19 CP-FIT **context variable** hypotheses [55] (Appendix 7.1.).

7.1.3.a. Organisation and Team Characteristics

CP-FIT hypotheses 24 to 31 suggest that having **resources**, strong **leadership support**, dedicated champions, and strong **teamwork**, as well as effective **intra-** and **extra-organisational networks,** are factors that contribute positively for the feedback to be taken seriously and to be integrated into organisations. A **workflow fit** for the tools and avoiding **competing priorities** in feedback further ensures smooth adoption into existing systems and workflow patterns, without distracting from essential duties [55].

Users mentioned resource constraints in interviews as an important factor that prevents them from using the tool and/or using feedback in their efforts to improve QoC. They also mentioned competing priorities as part of their time constraints, with many tasks at hand and staff shortage being a serious challenge. As a result, sometimes the feedback is only reviewed at home after work, something the users do voluntarily outside of their shift as they care about improving performance. There are some inconsistencies in leadership and organisational support as they are not able to provide the necessary resources to improve performance. However, at the same time the tools trigger more teamwork as people come together using feedback to advocate for more resources, although the success of these efforts are questionable. Teamwork and champions positively impact the use of the dashboards and the efforts to improve QoC. However, as suggested by some users, it could be important to evaluate what kind of leadership, champions and teamwork works best.

Furthermore, my data revealed a network of actors with different views on dashboard usage and maintenance responsibilities. This links to CP-FIT hypotheses 29 and 30 that say feedback is more effective when A&F interventions are integrated into organisations or teams with strong internal and external communication channels. Indeed, interpersonal and inter-organisational relationships play a key role in improving dashboard use, with users discussing performance with peers beyond their own facilities. The networks surrounding dashboard use could benefit from further investigation, e.g., using a social network analysis method, since the study showed that dashboards have more interaction points and underlying influences beyond the immediate use of the tool.

Both CIN and NEST implementers have a network of partner hospitals, universities and governmental support, and yet there are challenges around integrating feedback systems into the complex existing information systems and existing workflows (with DHIS2 and interoperability constraints). This complexity of the information systems underscores the need for better data management strategies and robust IT support to optimise the health systems in which the dashboards could fit. The implementers and users agree on that, and users value the CIN data system as a significant enhancement of the routine information system. Implementers also recognised that the effectiveness of their tools heavily depends on the level of organisational support and resources available. NEST implemented an entire bundle of interventions that users appreciate; however, "*remember we are in Kenya, there is never enough staff*" was a common refrain. Additionally, they see the importance of robust support systems within hospitals to address issues of staff turnover and resource constraints, but there is only so much they can do.

7.1.3.b. Patient Population

CP-FIT hypotheses 32 to 33 suggest that **choice alignment** ensures that feedback calculations exclude patients who refused aspects of the care that was being measured, and **clinical appropriateness** excludes patients whose care cannot be safely optimised further, both keeping feedback accurate and relevant [55]. My study could not find any choice alignment according to CP-FIT hypotheses. However, the interviews revealed users' concerns with the inconsistencies in source documents (patient files) and referrals of patients with low chances of survival, as it unfairly reflects on them as performing poorly. Implementers are aware of this issue, and CIN added information to the dashboard about how many of the admissions were referrals to acknowledge the challenge the facilities have with a high referral rate, which users seem to appreciate.

7.1.3.c. Co-Interventions

CP-FIT found that the feedback studies they included used 8 other quality improvement interventions besides the feedback, of which the following 4 seemed to impact feedback and to balance out the shortcomings of resource and health professional characteristics outlined above (hypotheses 34 to 37) [55]: **Peer discussion** encourage recipients to discuss their feedback with peers, which can also increase teamwork. **Problem solving** and **action planning** help users to identify issues and develop solutions for suboptimal performance, turning feedback into guidance for actionable improvements. **External change agents,** as in providing additional staff, help to support the implementation process. Both dashboards are perceived positively by clinicians for their potential to drive improvements in newborn care irrespective of co-interventions, due to the appreciation of having information ("*then we know*"). Co-interventions that CIN and NEST implemented came up in interviews as something users valued, e.g., NEST providing more equipment, and CIN providing a data system that can be trusted as well as creating a network and possibilities of peer-to-peer interactions during CIN meetings. However, the high workload and staff shortages suggest that without additional support, such as increased staffing or streamlined processes, the potential benefits of these co-interventions might never be fully realised. This highlights a potential need for targeted workforce co-interventions that address the specific challenges faced by healthcare teams and evaluate them alongside feedback interventions. For further problem-solving and action-planning interventions, both implementers could consider adjusting the current design and implementation of dashboards, e.g., interactive data visualisations, integrating performance tasks to the dashboard.

7.1.3.d. Implementation Process

Brown et al. also found that the implementation process plays a crucial role and impacts all feedback cycle processes with these specific requirements having a positive effect (hypotheses 38 – 41): Tailored feedback to the specific needs of the organisation and staff improved workflow fit, and hence **adaptability** of the feedback tool is an important aspect. **Training and support** on using the tool helps users to improve their skills in quality improvement, and with **observability**, meaning the tool demonstrates its benefits, people are more likely to engage with it. Furthermore, interventions perceived as expensive in terms of time, human or financial resources are less effective (**Cost**) [55].

My study findings highlight the technical and procedural challenges, such as the need for better training and support practices. However, users did not mention this as something that hinders effective implementation but rather in terms of interpreting the graphs, which is something that training indeed could address while a more user-friendly design could possibly be more cost-effective. Addressing these issues through tailoring training and visual feedback, creating more intuitive user interfaces, and streamlining processes, and providing user-specific support could upgrade the tools' usability and impact. The dashboards could be further improved if locally adapted to user (group) backgrounds and training needs. Adopting more interactive learning and teaching methods could support the effective use of dashboards and the visualisations they contain [74], [228].

Further, CP-FIT states that when feedback gives recipients a sense of **ownership**, rather than feedback being imposed, it is more successful, as it gives autonomy and motivation to improve care (hypothesis 42). However, Brown et al. also reported that they found limited support that including recipients' input to design and implementation at an early stage would be helpful [55], [229]. My user interview data suggests that while clinicians appreciate the feedback provided by the dashboards and they own the feedback (as in the content of the dashboard), they also see the implementers as the owners of the tools (as in the responsibility to improve them for use). In contrast to CP-FIT findings, users are convinced that if they were involved more in the design and implementation from the beginning, the dashboards could be even more successful.

7.1.4. Expanding CP-FIT with Design Principles for Newborn Dashboards in Kenya

Brown and colleagues acknowledge as part of hypotheses 14 and 15, that information on how to improve feedback towards user-friendly designs is limited and suggest reducing content or using charts [55]. Thus, incorporating broader design literature alongside CP-FIT in my study provides more user-specific insights for the two newborn care dashboards in Kenya.

7.1.4.a. Main User Groups and Interpersonal Relationships

My study included perspectives from implementers and examined the roles of primary and secondary users. Hospital staff are seen as primary users, but interviews highlighted that some of the hospital staff also view the implementers as the main users. Consequently, they expect implementers to handle data management, design, and training for effective tool use and see them as the dashboard managers. Their own role remains in using the dashboard content for improving data and care quality.

7.1.4.b. Design and Integration into Information Systems

Although CP-FIT is informed by pre-existing technology theories that include usability to some extent, it lacks integration of detailed design principles. CP-FIT focuses on feedback tailored to the recipients; however, it does not incorporate UCD principles. For example, it does not consider continuous, automated user feedback and iterative cycles to improve the feedback tool design. A study by Tsang et al. that used CP-FIT assessed electronic A&F systems and similarly found that usability was not largely part of the identified studies' feedback strategies. Out of 88 e-A&F systems and dashboards they examined, only 16 presented usability findings

[108]. My study expands on this by clearly showing the importance of testing usability as what users say and think differs from what they (can) do and emphasises the need for a design strategy to integrate users' perspectives that are essential for its effectiveness. Specifically, the interpretability of visualisations requires attention to ensure feedback is actionable. Simplified feedback visualisations, for example, using only simple bar charts or tables, and co-design are critical. Interviews suggest that mistrust of feedback can arise from both poor data quality and difficulty interpreting graphs.

7.1.4.c. Context

My survey offered limited insight into contextual factors affecting dashboard effectiveness, but the qualitative data highlighted the complexity of the environment. CP-FIT is based largely on studies from non-African and non-newborn care settings and does not specifically address complex health information systems. Addressing these will require a more context-tailored approach that goes beyond the CP-FIT framework. Most interviewees envisioned that the dashboards will be integrated into existing information systems, such as DHIS2, at some point. However, these data and information systems themselves must be supported by better data management strategies and IT support to become effective. In fact, many Kenyan hospitals have purchased limited forms of digital health record systems, but these vary in levels of interoperability, threatening options for using data for feedback. Muinga et al. further report that users have very similar issues with these systems: system usability, inadequate training, infrastructure and systems support that may result in distrust of data [230]. Systems challenges will further complicate the existing difficulties in interpreting graphs and using feedback to improve care when teams are under-staffed or face other resource constraints.

7.1.4.d. CP-FIT Fit for Purpose: A&F Dashboard Interventions in Kenya

CP-Fit is based on 65 papers evaluating 73 feedback interventions, focussing on qualitative assessments, with usability noted in one of the hypotheses. It draws on 30 pre-existing behaviour change theories and concepts and provides 42 hypotheses based on a variety of interventions. I found that CP-FIT provides a solid foundation for assessing dashboard feedback interventions and, as such, serves as a guide to evaluating how to deploy quality improvement dashboards. Its consideration of social dynamics, such as teamwork, networks, and performance improvement, and its use of underlying behaviour change theories make it a strong theory. All mechanisms that CP-FIT identified [55] – complexity, relative advantage, resource match, compatibility, credibility, social influence, and actionability – acted as key facilitators or barriers to the effectiveness of dashboards in my study too.

However, it may not fully capture the complexity of dashboard feedback systems for the Kenyan and SSA hospital setting, where feedback interventions are often implemented by research and philanthropic actors. In addition, it may overlook the importance of understanding specific user needs, capabilities, and interactions due to its lack of design principles that contribute to the theory. A&F interventions can be informed by applying the user empathy map by placing the (primary) user at the centre of the feedback design as recommended by user-centred design and design thinking principles, combined with an understanding of other (secondary) user perceptions (for example, implementers and other stakeholders).

My study furthermore highlighted the value of testing feedback interventions with qualitative and quantitative research methods, while CP-FIT is primarily based on qualitative studies. In fact, the quantitative analysis was crucial for challenging the findings from the Likert scale and qualitative data and showed the importance of testing accurate interpretation rather than relying on user perceptions only. It proves, in addition to the scoping review findings, that dashboard implementers should test usability quantitatively and report these findings more frequently in the future. Despite the important additions from the design literature, my findings show that CP-FIT can effectively be applied to identify many important variables and influences that help shape future design iterations.

7.3. Linking Findings from Empirical Study with Scoping Review

As seen from the scoping review (Chapter IV), newborn dashboards and scorecards have been implemented in various settings to monitor and enhance neonatal care, predominantly but not exclusively in HIC. They are said to be ‘valuable tools’ in improving clinical decision-making and patient care as they can lead to timely interventions and clinically relevant practice improvements [156] [155], [231]. Examples include dashboards in Canadian and UK hospitals that have been used to improve care by helping institutions to benchmark performance (with traffic-light systems), identify gaps, and use information for strategic improvements [160], [232].

Most publications reviewed do not extend to give guidance on testing that interpretations of dashboards are accurate and useable for action. Moreover, there does not seem to be a strategy published to date that provides an overview of how the A&F literature and design principles can be used together to optimise the accurate interpretation of feedback to support action. In fact, although most studies mentioned the importance of engaging the users in the development process and some recognised that thinking about users’ experiences is important, most studies do not report how they assessed the usability or interpretability of their dashboard tools. It is, however, a significant challenge. Each setting has unique requirements and a range of users, making a one-size-fits-all solution nearly impossible. The question arises whether it is feasible to develop generic design and implementation guidelines for newborn dashboards to improve care if they are required to be tailored to specific settings.

Furthermore, only limited evidence was found from LMIC settings, especially in Sub-Saharan Africa, where newborn mortality is high, and hospital resources are constrained. It is questionable whether evidence from other settings would be applicable to dashboard and feedback interventions there. The success of a newborn dashboard, for example, in Canada, depends on factors such as interdisciplinary collaboration, accountability, and alignment with organisational priorities [160]. However, implementing dashboards also requires significant investment in terms of financial and human resources to ensure data quality and system effectiveness [233]. As Flannigan et al. report, their dashboard in Northern Ireland helped to make a case for additional staff [148], which corresponds with my interview findings as people use the dashboard for resource advocacy; however, we have to "*remember we are in Kenya, where there is never enough staff*", which was mentioned as a challenge remaining beyond control. My work suggests that testing user experiences and capabilities for a specific setting with a mixed methods design is necessary when making feedback intervention investments, as what users say they can do differs from what they can actually do. This becomes most apparent when considering the quantitative testing of dashboard interpretation success.

7.4. General Discussion

In this section, I will present some of the challenges and paradoxes that became apparent in this study in reflection to the literature. I come back to the approach to translate data into action (see Figure 41) to reflect on each of these steps and how dashboard designs and implementation can be tailored to provide feedback and thus drive action for quality improvements in newborn care in Kenya.

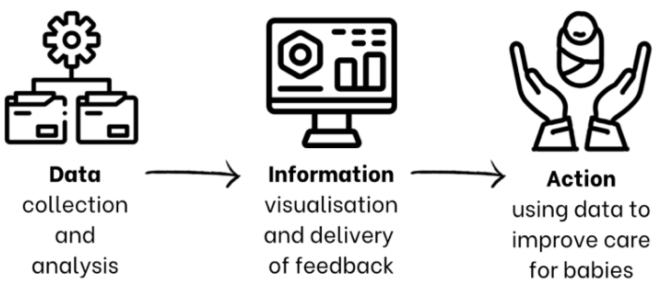


Figure 41: Reminder of data for action and information dashboards as a precursor for action.

7.4.1. Neonatal Data for Action

7.4.1.a. KPIs and Standards

KPIs are essential to monitor QoC and present newborn information in a structured way, with process indicators providing insights into the necessary steps that are important to take when improving outcomes [5], [38], [59]. CIN and NEST dashboards largely present relevant newborn KPIs (based on best practices and standards) that span care outcomes and processes that study participants agreed are important. Currently, the CIN feedback provides a larger amount of KPIs, while NEST limits feedback to one page. The variation of KPIs presented in dashboards seen from the scoping review underscores the need to customise indicator selection (dashboard content), balancing between global recommendations and locally tested indicators that are needed by the end users for quality improvements. NEST, for example, added resource availabilities (staff and equipment) to their dashboard, which users found very valuable for contextualising their care challenges, advocating for resources, and addressing constraints, suggesting that tracking these can improve QoC and inform broader health system investments.

When using the talk-through methods, users often focus on outcome KPIs, such as mortality (for NEST this could be supported by the prominence of showing mortality on the top), stating that the feedback from dashboards is useful because they "*now know the numbers*." When asked how these numbers help to improve care and outcome, users shift to the process indicators as essential information for guiding improvements. However, users often struggled to interpret them. Neither of the tools provides information on how to improve things. Users mentioned that they discuss feedback information during team meetings, and process indicators can be particularly sensitive to improving QoC and guiding action [5]. Yet, only 38.5% of users correctly interpreted the process indicators and care pathways in the NEST dashboard using grouped bar charts. In contrast, interpretation rates were higher for CIN's simpler bar chart visualisations of process indicators, suggesting the issue lies in the design rather than the indicators themselves. This raises concerns about the effectiveness of dashboards in driving action if users cannot fully understand data visuals.

7.4.1.b. Data (Sources) Rich, Information Poor?

KPI reporting depends heavily on the (quality of) data collected. In Kenya (and other LMICs) where neonatal mortality remains high, data systems are often fragmented, relying on a mix of paper and electronic records, which can prevent data collection from being consistent and of high quality and, therefore, hinder the full potential of data-driven improvements [15], [24]. Even in countries with advanced digital health data systems like Canada (e.g., the BORN Ontario network), newborn initiatives have created their own data and information systems to ensure data quality and specific newborn needs are addressed [155], [232]. CIN created a newborn data and information system for neonatal data that tries to ensure data quality [21], [22], [116], [234], and that has evolved over time to improve data quality by incorporating feedback from users during regular CIN meetings or hospital visits. NEST benefitted from CIN's system, showing the importance of collaboration rather than competition risking data silos. Both CIN and NEST work on the underlying data issues as a crucial part of the dashboards and ongoing quality improvement efforts. In creating these new systems, 'rich' data (extensive and of better quality than others) are now available in Kenyan NBUs. And yet, many users still report hesitance to trust the data and the paper-based information system it relies on. Despite this, many respondents express the need for enhanced data systems nationwide and the desire for faster implementation of digital systems and dashboards. To conclude, newborn data in Kenya is 'rich', as there are many collection efforts, but for data truly to be rich, it must be of high quality, and it depends on systems such as the CIN.

7.4.2. Designing Newborn Hospital Dashboards

7.4.2.a. Information Displays for Dashboard Usability

A dashboard primarily consists of data visualisations to translate data into concise, rapid-to-understand visuals, making the layout and interface critical components alongside other dashboard user experience aspects. When asking users about dashboard usability and graph interpretability, the responses "*it is very easy to use*" and "*very useful*" were common. Effective data visualisations require careful selection of graph types to ensure that data is easy to interpret as otherwise, information cannot be actionable, or worse, distort the underlying data. Human cognitive psychology research on which graph types are easier to understand has been extensive, providing guidance on how to translate data into information and inform best practices on dashboard visualisations [25], [74], [93], [98], [100]. Tools like data vocabularies, data storytelling books, and a chart chooser can help guide when selecting graph types [74], [96], [97], [98], [235]. Figure 42 shows that no chart is inherently wrong, it is rather about what the aim is when translating data into information.

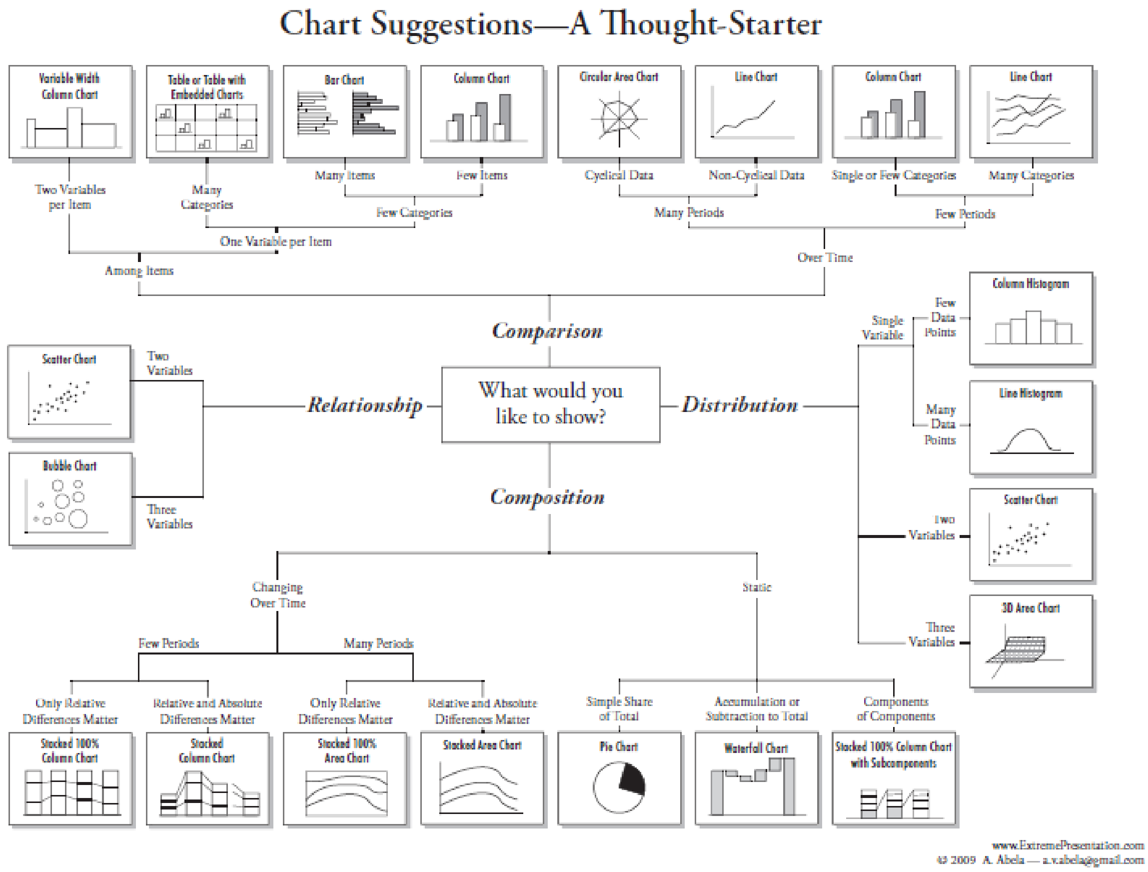


Figure 42: Data Visualisation Chart Chooser by © Andrew Abela, 2010 [96], [235].

Cognitive science shows that humans can process only limited information at once. A good dashboard layout does not require the user to calculate or take more than a couple of minutes to understand the information presented [93], [98], [100]. When a graph is selected, simplicity in data visualisations is known to be crucial for people to understand the information they try to communicate. Simple graphs, supported by Gestalt principles (using colour, spatial positioning or hierarchy, and motion), facilitate quick interpretation and encoding of data [25], [184]. In terms of the colour, red was highlighted by users as helpful to identify quickly where action is required, and the purpose of dashboard traffic light systems seems to work. However, users also mentioned the challenge of still relying on print versions, which are only possible as black-white prints due to resource constraints, when wanting to move information to other places. Although traffic lights can be useful, in general, it is recommended to avoid relying entirely on colour-coding, particularly red-green combinations, and instead use additional visual cues like shapes or text to ensure accessibility for colour-blind users [25], [74].

Testing graph types is shown to be one of the most important parts of this study's assessment of CIN and NEST dashboards. More careful selection and testing of graph types is essential to align visual designs with user needs and skills for the specific context of often-hectic NBU and hospital environments in Kenya. Testing different visual representations in context across all user types, minimising cognitive load, and creating intuitive dashboard layouts can reduce the need for training, which is currently an additional burden for senior nurses due to high staff turnover and the responsibility to guide incoming staff. User testing identified some graphs requiring urgent changes as their interpretability is below 40% of participants being able to understand them, and users often overestimate their interpretation skills (see Table 8). Hence, in future iterations, developers should actively test dashboard user and interpretation skills in addition to listening to users. As dashboard designs may never fit all user group needs, testing can also provide insights into specific user group skills, providing opportunities to tailor training needs alongside design adjustments.

7.4.2.b. User-Centred Design: One Does not Simply Design a Dashboard

When developing a dashboard, several key aspects from the UCD literature should be considered as guiding principles. The UCD process includes understanding users and their needs through observation, empathy, and ideation, followed by prototyping and iterative testing (see Figure 9). CIN and NEST have developed their dashboards based on some of these principles, though their approaches differ, and this was the first study testing the visual content and exploring the user experience and all sub-elements (see Figure 7).

The first step of UCD is empathising with users to ensure the dashboard aligns with their needs, goals, and workflows [89]. CIN and NEST implementers have conducted health facility assessments (users' environment) and consulted users throughout the process (users' desires). However, users and some implementers were not entirely certain whether the dashboards were co-designed or whether feedback was obtained rather informally, unstructured, and introduced later in the process. The **user empathy** map (see Figure 43) is considered the foundation of understanding the user, what they say, think, feel, and do, which has helped in my study to identify the discrepancy between users saying overconfidently they can interpret while most failed to do so, which would not have been visible through verbal user feedback alone.

Figure 43: User empathy mapping exercise completed with findings from quantitative and qualitative data.

The **iterative** nature as a principle of UCD is reflected in the evolution of both dashboards. **Continuous** feedback has already led to improvements in their display (e.g., NEST changing from a pie to bar chart) during the time of this study. However, as seen from the results and discussion above, there is still room for improvement in current dashboard displays. When comparing CIN and NEST, both dashboards have the same goal (improving newborn care), but they diverge in their design, and users have clear visions of how they could further complement each other as tools that both have their legitimacy ('they are talking different things'). Users suggested a real-time data-driven, condensed one-page dashboard like NEST to show trends and prioritisation for immediate action-oriented tasks, while CIN continues to provide a more comprehensive set of indicators as monthly/quarterly feedback to support quality improvement team meetings and strategy planning.

Research shows that 3-5 iterations are at least needed to resolve major usability issues, and ongoing updates post-launch are required to adapt to real-world changes [236]. My study completed the first design process round, testing 'usability' and adding comprehensive user experience information to empathise with users and brainstorm and ideate the next dashboard version. CIN and NEST developed somewhat effective dashboard prototypes that now have the chance to evolve through iterations and re-design. These iterations should involve users and balancing users' and implementers' knowledge for sustained improvements.

7.4.3. Implementing Dashboards in Kenya: One Size Does Not Fit All

7.4.3.a. Scaling and Integrating Dashboards into Systems and Environment

With strong findings on the contextual challenges, it is evident that successful dashboard development requires to evolve beyond the technical features of the tool itself and to integrate into existing systems. Implementers acknowledged this as an important but probably the most difficult next step.

Interoperability is fundamental: dashboards should integrate seamlessly with existing health information systems, just like other (digital) health information tools [111], [230]. Without this, high-quality data, information and knowledge sharing may remain isolated in silos, limiting the broader impact on QoC. As described earlier, dashboard design is not a one-time event; it must be iterative and continuous. As health systems evolve, so do user needs, the environment they work in and the technical infrastructures. Continuous and adaptive refinement is, therefore, a necessity for any electronic information system, including dashboards, to remain functional and relevant. This requires the capacity to regularly update the dashboard's features to align with changing HIS requirements and emerging user needs. At the same time, it requires to be **adaptive** to the real-world contexts in which healthcare providers operate. Understanding how dashboards fit into daily workflows is essential for their uptake and usability. Dashboards should reduce, not increase, the complexity of users' tasks by fitting naturally into existing routines and systems. **Scaling** newborn hospital dashboards, just as other digital health interventions, to broader healthcare systems is a complex task that will require balancing standardisation and flexibility [111], [112]. While standards ensure data consistency and interoperability, local contexts and changing environments may require adaptability in design for further scaling CIN and NEST.

7.4.3.b. Learning, Teaching and Communication Opportunities

Dashboard users often asked for more training, so the question arises what kind of training is needed. Taking the interpretation difficulties as one of the key learnings from this study, I assessed what effective methods for teaching data visualisation interpretations as this may lead to better design and implementation of dashboards, including the next CIN and NEST design iterations.

Step 1: Active learning strategies as part of dashboard design:

Scrolly-telling is an interactive, immersive digital storytelling technique where the narrative unfolds as the user actively scrolls through feedback controlled by the user themselves [74], [237]. Interactive step-by-step guides, walk-through video tutorials, and 'scrolly telling' tutorials can improve user performance, especially for unfamiliar visualisation types and are generally preferred by users [238]. This would be easiest once dashboards are fully digitised, and a training programme could be implemented. **Interactive data facts/** information visualisations can be incorporated into systems to help users interpret visualisations correctly and explore alternative visualisations and add-ons to communicate data differently [239]. This could particularly help with the mistrust users mentioned when not being able to interpret specific visualisations.

Step 2: Teaching methods as co-intervention of implementation:

Teaching methods can engage participants directly, for example, with use cases and real data visualisations (e.g., from past performance), where learners are firstly helped to understand the overarching concepts of data interpretation before secondly diving into the specific case. This is claimed to make complex information more digestible for learners [228]. Educational psychology principles, such as cognitive load theory, can help to ensure information is presented in manageable chunks, preventing learners from becoming overwhelmed [104]. These learning styles and methods, and creating *"formal"* learning materials, would require the implementers to see what works best in a specific setting and for different user groups with different capabilities.

7.4.4. Overall Strengths and Limitations

My study successfully addressed the academic and practical gap identified in Section 2.4. I could bridge the academic and practical need to assess data use and dashboards as solutions to help healthcare workers in improving QoC for newborns in Kenya, and for health systems actors to track progress toward SDG 3.2. The study shifted the focus from data and KPIs to understanding users' needs and moved from 'what' content dashboards should contain to exploring 'how' they should be designed through systematic practical testing and exploring user perspectives.

General limitations related to the research methods were discussed in each chapter. The scoping review revealed a need for more evidence on how newborn dashboards were designed and tested, in particular for settings where newborn mortality is high. My study closed this gap by providing evidence in this area, and with a mixed methods design, combining qualitative and quantitative data insights, providing a comprehensive understanding of A&F dashboards in Kenya.

The strength lies in the theoretical use of CP-FIT, which I extended by integrating design principles ensuring my survey and interview tools were grounded in theory and practicality. The discrepancies between users' self-reported perceptions and actual performance in interpreting dashboard content, confirmed the importance of mixed methods, combining user feedback with real-world testing. The aim was to include users and potential users of CIN and NEST (including those who are unfamiliar with the tools). However, most participants were familiar with one of the dashboards. Including more 'non-users' could have possibly identified challenges that current users have already learnt to mitigate the[240]. A strength was to include dashboard developers, who contributed with valuable views on technical feasibilities and diverse (user) challenges spanning across health systems and hospital levels. Interviews with both, implementers and users, provided a comprehensive understanding of dashboard developments, which extends typical qualitative A&F and UCD studies that often focus on feedback recipients primarily. For future studies of (digital) dashboards in the newborn care context in LMICs, a multi-user experience (MUX), multi-user centred design (MCD) should be considered [240], as this approach will help to include other users' perspectives.

Observations showed survey participants needed more time to answer multiple-choice questions when interpreting graphs, and interviewees struggled to immediately explain how they act on information. While my study was limited to a talk-through method due to the COVID-19 pandemic restrictions, I suggest future studies will use systematic qualitative walk-throughs and observations. Alternatively, a (un-)moderated remote testing and walk-through of the dashboards can be conducted to observe people's interactions in real-time, once dashboard prototypes are fully digitised [241].

The focus on Kenya could restrict the generalisability to other settings where NEST has implemented their dashboard. However, it still serves as a foundation and important first step in testing newborn dashboards in Sub-Saharan Africa. Other limitations, like in any research, include potential biases that are difficult to eliminate, and these were reported for the qualitative and quantitative components earlier. However, a strength of this study is regular sense-checking and feedback from the research community, as I presented at multiple conferences and meetings (Appendix 7.2.)

7.5. Future Work & Research

Dashboards can offer a powerful way to make use of data for decision-making. However, if not well-designed and if not based on quality data, they can also

- make data harder rather than easier to understand and complicate interpretation, and/or cause more workload for healthcare workers,
- risk the misuse of data for decision-making, which in the worst case can possibly even cause harm to the patient, and
- misrepresent realities and narratives from the ground to the national and international community, and potentially harm organisational goals and reputation.

My study offers key insights that help to minimise these risks for newborn hospital dashboards in Kenya. However, it is important to acknowledge that some areas remained beyond the scope of this thesis and are worth further investigation, which all point towards a multidisciplinary approach requirement:

Re-design and future iterations

While I could complete one foundational iteration, future work should prioritise testing new dashboard adjustments and redesigns. New designs and testing could include spatial positioning of KPIs (process/outcome/resources), colour schemes, interactive visuals and descriptions. In future studies, user process mapping with walk-through interviews and eye-time-tracking when using the dashboard could be conducted, which could add further depth to user experience and usability testing. A time series approach could monitor improvements through multiple iterations, as seen from the BORN study [156].

Dashboard co-interventions

Co-interventions could be further explored, such as for example extending the CIN and NEST dashboard content to other indicators and feedback strategies:

- **Maternity care:** Many factors contributing to newborn mortality are rooted in maternal health [12]. Extending the dashboards and include maternal indicators as a collaborative effort of both wards could provide more insights and improve outcomes. NEST considered this already.
- **Teamwork:** Users mention the importance of teamwork (within their team and across departments), such as neonatal and maternity units. Indicators that track team collaboration and work culture, such as interdisciplinary communication and coordination, could track where hospitals perform well in these efforts. However, this could also be conducted as a study using social network analysis methods.
- **Individual feedback:** Individuals receiving feedback can also help to encourage staff to reflect on their own role and purpose in improving care outcomes when working as a team. Incorporating this into team-based assessments could help managers to manage human resources.
- **Peer and colleague feedback** on professional practices, could promote accountability and shared learning as seen from other feedback interventions providing best practice guidance, for example for antibiotic prescribing [242]. Encouraging open dialogue and feedback among healthcare workers could improve communication and allow staff to learn from each other's performances.
- **Patient feedback:** Ultimately, the most important beneficiary of quality data and care is the patient, the newborn. One critical gap in the current system is the lack of **patient feedback**, which for newborns would be the mother. Mother's feedback and views on improving the quality of maternal and newborn care have shown to provide valuable inputs, e.g., for improvements in counselling and support on breastfeeding during the neonatal phase and in communication between mothers and health workers who care for these babies [243], [244], [245].

Dashboard impact on outcomes & cost-effectiveness analysis

The effect on clinical outcomes can and should probably be measured. Coupling dashboard use with other interventions could be tested using a Randomised Controlled Trial (RCT) to measure direct impacts on healthcare outcomes [246]. However, conducting an economic evaluation of A&F dashboard intervention could probably reveal valuable insights for scaling and sustainability plans of these interventions.

Emerging technologies

Existing and emerging technologies used by hospital staff (e.g., smartphones, artificial intelligence applications) will probably require future work for dashboard development considerations. While technology can improve the QoC, it may also distract healthcare workers, introducing more complexities and unintended consequences. The finding that users mostly use smartphone, even during work, although unclear how much of it is 'for work' or 'during work', and whether these are contributing or distracting from tasks at hand, indicates there could be potential benefits and downsides the require consideration in future dashboard iterations.

Joy in work

The burden of a high workload during staff shortages, mentioned by many users, suggests that future studies should explore workplace well-being in relation to performance interventions. The *IHI framework* for improving joy in work provides guidance on how leaders can improve joy at work, which contributes to better performance [247]. However, there is no evidence on how this applies to a low-resource hospital setting. Although users in my study showed accountability, there is also a psychological drain that was mentioned as *"there is only so much one can do".*

Scaling to address SDG 3.2.

Other countries where neonatal mortality remains high would probably benefit from similar studies for contextual insights. My study can serve as a foundation for this work. With NEST already operational in other African nations, there is an opportunity for a cross-country QoC analysis for dashboard interventions contributing to achieve SDG 3.2.

Post-2030 era

Finally, future research must address how newborn data dashboards integrate into national and global health infrastructure when the SDGs are re-evaluated soon. Future research should analyse the broader data ecosystem and the role of metadata (dashboards). A new instrument to test public or government dashboard usefulness was published after I conducted the survey, which could be used for future testing [248]. Dashboards like the WHO's maternal and newborn health portal will probably play a key role in supporting high-level decision-making. However, these tools rely on national data and need to ensure transparent, reliable data across all levels of the health system. Otherwise, users might face the same data quality trust issues as the participants from this study.

VIII. Conclusions

8.1. Recommendations Drawn from Study Findings

Based on my study's findings and the reviewed literature, I present three key recommendations for consideration when developing dashboards for quality improvement, drawing on the case of newborn care in Kenyan hospitals as an example.

8.1.1. Foundations for Developing Quality Improvement Dashboards

The development of quality improvement dashboard interventions (and this is probably applicable to any setting) requires an evidence-based foundation to reduce avoidable inefficiencies.

As Brown et al. suggests, feedback, and thus A&F dashboards, are most effective when they offer actionable insights, meaning clinicians can understand the impact of their actions, which will then encourage quality improvement [55]. CP-FIT, a theory-informed approach that incorporates behaviour change frameworks, can indeed serve as a foundational part when designing dashboards, as these theories help to shape feedback in ways that motivate behaviour change among healthcare providers [56]. However, A&F interventions require a stronger use of best practices derived from well-established information science and design literature [25], [74], [98]. UX and UCD principles are essential for developing information tools that fit seamlessly into clinical workflows and serve diverse user needs [26]. Dashboard developers must focus on clarity and simplicity, minimising cognitive load, and ensure underlying data quality to prevent misinterpretation. Insights from cognitive load theory and visual design principles, such as Gestalt, should guide the design to ensure that dashboards are intuitive and reduce unnecessary cognitive burden [98], [100]. Simplifying visualisations and refining visual elements, such as replacing ineffective pie charts and tailoring elements like traffic light systems to user needs can significantly improve usability and interpretability [74]. Consistently driving towards the most user-friendly data visualisations as part of dashboard designs includes a) to empathise with the user and b) to test usability and interpretability of data visuals [25], [26], [76], [89].

A successful strategy involves engaging multiple user groups, testing prototypes, and using these assessments to create tools that are both theoretically sound and practically useful. Effective feedback should be a two-way process, where user feedback is integrated to inform developers. Empathy mapping is a crucial first step, followed by an iterative design strategy. Placing the user at the centre of dashboard design and implementation adds the necessary contextual nuances to adapt evidence from other settings effectively. These practical lessons can then guide the first iteration. At the same time, balancing implementer and user knowledge is valuable. Implementers may also act as users (e.g., tracking progress) and can provide insights that balance user needs with technical or operational experiences, leading to a more robust dashboard design.

8.1.2. Design Iteration and Sustainability Strategy

Continuous design iterations – at least three – are recommended to improve the user experience and interface of a product, service or system, and to ensure long-term effectiveness [236]. Continuous simplification and including user needs should be central for every dashboard iteration [3], [85], [100]. As Don Norman once said, "*Good design is actually a lot harder to notice than poor design, in part because good designs fit our needs so well that the design is invisible, serving us without drawing attention to itself*" [26].

Based on the findings of this study as the first iteration, CIN and NEST dashboards have the evidence to improve visual design while accounting for external factors that influence the effectiveness of their dashboards. Since the start of this study, CIN and NEST dashboards have evolved. However, future iterations will remain critical particularly to adapt to changing users as well as changing needs, and to accommodate the scaling of dashboards across different health system levels. As dashboards expand, they will have to address diverse needs. While hospital staff at the micro level may require real-time data for patient care, meso- and macro-level managers and government officials may request aggregated data for strategic decision-making [249]. Ensuring adaptability and tailored content will improve usability and effectiveness in improving neonatal care. The design process, refinement, and implementation should be continuous rather than a one-time project to ensure dashboards remain suitable, adaptable, and sustainable in a changing environment.

Co-interventions, such as a ‘customer service’ and targeted ‘formal training’ (e.g. for data & information literacy), were recommended by users to improve dashboard use. Expanding to include maternal health indicators and patient feedback would address current gaps and create a more comprehensive quality improvement tool.

Training can provide hospital staff with the necessary skills to effectively interpret and act on data. However, innovative training that is integrated to software and supports self-learning can also enhance the usability and accessibility of dashboards, and evidence-based ideas were presented in this study.

The need for context-specific adaptation is especially pronounced in resource-limited settings, where the effective integration of dashboard technology must be carefully managed to prevent additional complexities. Funding bodies and implementers must ensure that dashboards streamline rather than complicate clinical workflows.

8.1.3. Unlocking Data 4 Action for Learning Health Systems

The concept of Learning Health Systems (LHS) involves integrating data and evidence into healthcare practices, operating through an iterative cycle of data collection analysis and implementation, guided by learnings and knowledge [250], [251]. As LHS aim to improve health outcomes, quality improvement dashboards play an important role to integrate real-time data analytics into healthcare practices and contribute to LHS [71].

To contribute to LHS, dashboards need to be built on robust and high-quality data ecosystems. This study showed that the underlying data and systems and the integration of dashboards is one of the biggest challenges implementers have. Evidence from HIC suggests that also data analytics capabilities in health systems are still underutilised [252]. The transitioning from paper-based to fully digital health records in LMICs, as well as feedback systems for QoC offers an opportunity for data-driven healthcare improvement.

However, if the underlying data quality and systems in place are an issue, users are unlikely to trust the feedback provided, and dashboards will never unfold their full potential. Thus, although it was important to shift focus from data to users, there is still a need to address the underlying data challenges. And dashboards will never be able to contribute fully to QoC improvements. These inadequacies in data collection and information systems lead to over-reliance on inconsistent, complex data, introducing uncertainty and mistrust into decision-making processes.

Furthermore, hospital staff wish to be connected, and their interpersonal relationships and teamwork matter when driving towards QoC improvements. The CIN and NEST dashboard initiatives have demonstrated some first successes in creating collaborations with and between partner hospitals. However, there remains considerable potential to further expand these networks, e.g., by using dashboards to connect different hospital teams. Users have expressed the desire for better connectivity across facilities with stronger platforms for shared learning and problem-solving.

Ultimately, effective dashboards depend on high-quality data, clear visualisations, training, and supportive learning communities. When these elements come together, dashboards become powerful tools for driving continuous quality improvement, informed decision-making, and ultimately better health outcomes, particularly for vulnerable populations like newborns in resource-limited healthcare environments.

8.2. Closing Remarks

Newborns admitted at hospitals and newborn units, receive complex treatments at the most vulnerable stage in life, susceptible to harm and at risk to die [253]. Effective quality improvement requires information that reflects which area require attention by healthcare workers to improve their performance. The WHO-UNICEF Expert and Country Consultation on Small and/or Sick Newborn Care found that many countries are not on track to achieving SDG 3.2. by 2030, and the challenges and reasons have been outlined by several sources, which identified weak health systems, inconsistent data systems, limited focus on neonatal care interventions, and scalability issues of interventions due to resource constraints [143]. Although the dashboards helped to improve data quality, there remain significant gaps in resources that dashboards cannot entirely resolve. Addressing these challenges will be crucial for achieving SDG 3.2 targets.

CIN and NEST are two initiatives that stand out in their attempt to use data to provide information and drive action in newborn care aiming to achieve SDG 3.2. With this study, there is now some evidence of QoC dashboard developments, so they can better contribute to these efforts for Kenya. Since the evidence is missing for other LMICs, there remains a need to focus on the specific factors relevant to each environment when designing a dashboard intervention to make use of data and for data-driven quality improvement. However, the foundations for designing a dashboard have been outlined above.

One major conclusion of this study is that we should not make assumptions about user needs even when they tell us dashboards are very useful and easy to use. Simply asking users what they want and like cannot replace researching what their needs are. More human- / user-centred design approaches in healthcare interventions that explore unspoken user experiences and unmet needs, understanding user behaviour, motivations, and long-term goals, rather than merely reacting to immediate feedback might provide better tools that serve healthcare workers for the benefit of their patients.

This thesis underscores the relevance of evaluating A&F dashboards for quality improvement in healthcare, since, as Don Norman's stated, *"poor feedback can be worse than no feedback at all, because it is distracting, uninformative, and in many cases irritating..."* [76]. It is a necessity to develop well-designed, meaningful performance feedback dashboards to truly drive quality of care improvements, which requires a blend of theoretical knowledge, practical user-centred design, and iterative improvement process strategies.

IX. Appendices

Chapter III Methodology Appendices

Appendix 3.1. List of CIN and NEST Partner Hospitals in Kenya

CIN (n=24) and NEST (n=13) partner hospitals ordered by counties (2022/23), counties in bold were part of the network during data collection for my study [22], [116] .

Counties	HOSPITALS	Collaborating with...
1. Bungoma	1. Bungoma County Referral Hospital	CIN & NEST
2. Busia	2. Busia County Referral Hospital	CIN
3. Embu	3. Embu Level 5 Teaching & Referral Hospital	CIN & NEST
4. Homa Bay	4. Homa Bay County Teaching & Referral Hospital	CIN
5. Kakamega	5. Kakamega County General Teaching & Referral Hospital	CIN & NEST
6. Kiambu	6. Kiambu Level 5 Hospital	CIN & NEST
	7. Thika Level 5 Hospital	CIN & NEST
7. Kirinyaga	8. Kerugoya County Hospital	CIN & NEST
8. Kisii	9. Kisii County Hospital	CIN
9. Kisumu	10. Jaramogi Oginga Odinga Teaching & Referral Hospital	CIN & NEST
	11. Kisumu County & Referral Hospital	CIN
10. Machakos	12. Machakos Level 5 Hospital	CIN & NEST
11. Migori	13. Migori County Hospital	CIN
12. Nairobi	14. Kenyatta National Hospital,	CIN & NEST
	15. Mama Lucy Kibaki Hospital,	CIN & NEST
	16. Mbagathi County Hospital,	CIN
	17. Pumwani Maternity Hospital	CIN & NEST
	18. Ruaraka Uhai Neema Hospital in Nairobi (RUNH), run by the Italian non-profit organisation Word Friends	
13. Nakuru	19. Naivasha County Referral Hospital,	CIN
	20. Nakuru Level 5 Hospital	CIN & NEST
14. Nyeri	21. Karatina Sub-County Hospital,	CIN
	22. Nyeri County Referral Hospital	CIN & NEST
15. Trans Nzoia	23. Kitale County Referral Hospital	CIN
16. Vihiga	24. Vihiga County Referral Hospital	CIN

Appendix 3.2. Confirmation Research Ethics Training

fhi360
THE SCIENCE OF IMPROVING LIVES

FHI 360

certifies that

Christiane Hagel

has completed the

RESEARCH ETHICS TRAINING CURRICULUM

March 16, 2020

Appendix 3.3. Ethics Approvals

3.3.1. Oxford Tropical Research Ethics Committee (OxTREC)Approval

Oxford Tropical Research Ethics Committee

University of Oxford
Research Services, University Offices
Wellington Square, Oxford OX1 2JD
Tel. +44 (0)1865 (2)82106
E-mail: oxtrec@admin.ox.ac.uk

UNIVERSITY OF OXFORD

Christiane Hagel
Nuffield Department of Medicine
Centre for Tropical Medicine and Global Health
University of Oxford

23 June 2021

Dear Christiane

Full Title of Study: A study to learn how to design and implement feedback tools/dashboards for improving quality of care for newborns in LMICs

OxTREC Reference: 532-21

Thank you for your email and for your minimal risk application form.

I am pleased to confirm that approval has now been granted for this study. This is valid for the first five years.

The documents approved for this study are as follows:

Documents:	Version:	Date:
Minimal risk form	02	21/6/21
Appendix -- PIS and consent	02	21/6/21

Any subsequent changes to the application must be submitted to the Committee as an Amendment. This should include a letter to give the reasons for the proposed modifications and all revised documents with changes tracked.

Please ensure that you submit a completed Annual Report form on every anniversary of this approval and a final End of Study Report. The relevant forms can be found on the OxTREC website: https://researchsupport.admin.ox.ac.uk/governance/ethics/apply/oxtrec.

Finally, please note the following **important information**:

Data safety—all studies
It is the responsibility of the PI to ensure that all data collected during the course of the study is stored and transferred safely and securely. Further guidance and advice is available from the Research Data Team.

Only studies that will involve storing human tissue samples in Oxford
If you are planning to import the samples into England, you will need to make arrangements before the samples are transferred to store them under the governance of a Human Tissue Authority (HTA) licence. It is a legal requirement that any tissue or fluid made up of or

Tel: +44 (0)1865 (2)82106
Email: oxtrec@admin.ox.ac.uk
Web: https://researchsupport.admin.ox.ac.uk/governance/ethics

containing human cells to be used for the purpose of research is stored on premises licensed by the HTA unless covered by an exemption. OxTREC approval is not a recognised exemption. Further information may be found on the University's human tissue governance web pages: https://researchsupport.admin.ox.ac.uk/governance/human-tissue.

Yours sincerely

Dr Rebecca Bryant
Research Ethics Manager, OxTREC

3.3.2. Kenya Medical Research Institute (KEMRI) Approval 3459

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840-00200, NAIROBI, Kenya
Tel: (254) 2722541, 2713349, 0722-205901,0733-400003, Fax: (254) (020) 2720030
Email: director@kemri.org, info@kemri.org, Website. www.kemri.org

KEMRI/RES/7/3/1 **May 07, 2020**

TO: **PROF. MIKE ENGLISH**
PRINCIPAL INVESTIGATOR.

THROUGH: **THE DIRECTOR, CGMR-C**
KILIFI

Dear Sir,

RE: **SERU PROTOCOL NO. 3459 (*REQUEST FOR ANNUAL RENEWAL*): A CLINICAL INFORMATION NETWORK- A TECHNICAL COLLABORATION WITH THE MINISTRY OF HEALTH AND COUNTY HOSPITALS TO SUPPORT AND IMPROVE STRATEGIES FOR AUDIT AND HEALTH SERVICE EVALUATION**

Thank you for the continuing review report for the period **May 11, 2019** to **March 23, 2020**.

This is to inform you that the Expedited Review Team of the KEMRI Scientific and Ethics Review Unit (SERU) was of the informed opinion that the progress made during the reported period is satisfactory. The study has therefore been **granted approval**.

This approval is valid from **May 11, 2020** through to **May 10, 2021**. Please note that authorization to conduct this study will automatically expire on **May 10, 2021**. If you plan to continue with data collection or analysis beyond this date please submit an application for continuing approval to the SERU by **March 29, 2021**.

You are required to submit any proposed changes to this study to the SERU for review and the changes should not be initiated until written approval from the SERU is received. Please note that any unanticipated problems resulting from the implementation of this study should be brought to the attention of the SERU and you should advise them when the study is completed or discontinued.

You may continue with the study.

Yours faithfully,

ENOCK KEBENEI,
THE ACTING HEAD,
KEMRI SCIENTIFIC AND ETHICS REVIEW UNIT.

In Search of Better Health

3.3.3. Kenya Medical Research Institute (KEMRI) Approval 4203

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840-00200, NAIROBI, Kenya
Tel: (254) 2722541, 2713349, 0722-205901,0733-400003, Fax: (254) (020) 2720030
Email: director@kemri.org, info@kemri.org, Website. www.kemri.org

KEMRI/RES/7/3/1 **May 27, 2021**

TO: **DR. JACKSON MICHUKI MAINA,**
PRINCIPAL INVESTIGATOR

THROUGH: **THE DEPUTY DIRECTOR, CGMR-C,**
NAIROBI.

Dear Sir,

RE: **PROTOCOL NO. KEMRI/SERU/CGMR-C/229/4203 *(RESUBMISSION OF INITIAL SUBMISSION)*: EVALUATING THE EFFECTS OF TECHNOLOGY AND WORKFORCE ENHANCEMENT TO SUPPORT NEONATAL HOSPITAL CARE IN KENYA.**

Reference is made to your letter dated May 19, 2021. The KEMRI Scientific and Ethics Review Unit (SERU) acknowledge receipt of the revised study documents on May 20, 2021.

This is to inform you that the Committee notes that the issues raised during the 310th Committee C meeting of the KEMRI Scientific and Ethics Review Unit (SERU) held on **April 29, 2021** have been adequately addressed.

Consequently, the study is granted approval for implementation effective this day, **May 27, 2021** for a period of **one (1) year**. Please note that authorization to conduct this study will automatically expire on **May 26, 2022.** If you plan to continue with data collection or analysis beyond this date, please submit an application for continuation approval to SERU by **April 14, 2022.**

Please note that only approved documents including (informed consents, study instruments, Material Transfer Agreement) will be used. You are required to submit any proposed changes to this study to SERU for review and the changes should not be initiated until written approval from SERU is received. Any unanticipated problems resulting from the implementation of this study should be brought to the attention of SERU and you should advise SERU when the study is completed or discontinued.

Prior to commencing your study, you will be expected to obtain a research license from National Commission for Science, Technology and Innovation (NACOSTI) https://oris.nacosti.go.ke and also obtain other clearances needed.

Yours faithfully,

ENOCK KEBENEI,
THE ACTING HEAD,
KEMRI SCIENTIFIC AND ETHICS REVIEW UNIT.

In Search of Better Health

Chapter IV Scoping Review Appendices

Appendix 4.1. Search Strategy

First search: 31/08/2020

All searches were adapted to meet the specific requirements of each database and were restricted to publications from (and including) year 01-01-2010 to 31-12-2020. Additionally, only articles published in English were included. With this search strategy and collating all results, a total of 9,073 publications could be identified. In the Cochrane Library, 1,989 trials were initially identified; however, these trials were not exportable and thus could not be included in the total count of screened articles. I manually screened the first 150 trials, sorted by relevance, directly on the website, and found none to be eligible for inclusion. Consequently, these were not represented in the PRISMA-ScR diagram.

Second search: 06/10/2024

Additionally, I assessed all databases using the same search strategy in October 2024 to determine the number of publications available from the period 01-01-2021 to 31-12-2023 for a potential update. However, I did not proceed with conducting the actual update. A total of 5,680 publications were identified (see results for each database in *italics* in table below), along with an additional 1,186 trials from the Cochrane Library. Following a recent update, the Cochrane Library now permits the export of all trial data, which will facilitate a more comprehensive review process in the future.

Database	Search Strategy	Results
PubMed	(((((((Neonatal[Title/Abstract] OR Newborn[Title/Abstract] OR new-born[Title/Abstract] OR Neonate*[Title/Abstract] OR Infant*[Title/Abstract] OR Premature[Title/Abstract] OR Preterm[Title/Abstract] OR Birth[Title/Abstract] OR Childbirth[Title/Abstract] OR Delivery[Title/Abstract] OR Labor[Title/Abstract] OR Labour[Title/Abstract] OR Natal[Title/Abstract] OR Postnatal[Title/Abstract] OR Perinatal[Title/Abstract] OR Prenatal[Title/Abstract] OR Partum[Title/Abstract] OR Postpartum[Title/Abstract] OR Peripartum[Title/Abstract] OR Intrapartum[Title/Abstract] OR Prepartum[Title/Abstract] OR Antepartum[Title/Abstract]) OR (((("Infant, Newborn"[Mesh]) OR "Premature Birth"[Mesh]) OR "Labor, Obstetric"[Mesh]) OR "Prenatal Care"[Mesh]))) OR ("Maternal-Child Health Centers"[Mesh])) **AND** ((((dashboard*[Title/Abstract] OR "score card*"[Title/Abstract] OR scorecard*[Title/Abstract] OR "score board*"[Title/Abstract] OR scoreboard*[Title/Abstract] OR	**3,102** *Update:* *1,713*

	graph*[Title/Abstract]) OR "visual display*"[Title/Abstract] OR visualisation*[Title/Abstract] OR visualization*[Title/Abstract])))) **AND** (((quality[Text Word] OR performance[Text Word] OR improv*[Text Word]) OR ("Quality of Health Care"[Mesh])) OR ((indicator*[Text Word] OR measure*[Text Word]))) **NOT (**graphene)	
Scopus *Search 1*	**TITLE(**Neonatal OR Newborn OR new-born OR Neonate* OR Infant* OR Premature OR Preterm OR Birth OR Childbirth OR Delivery OR Labor OR Labour OR Natal OR Postnatal OR Perinatal OR Prenatal OR Partum OR Postpartum OR Peripartum OR Intrapartum OR Prepartum OR Antepartum OR (Premature Birth) OR (Prenatal Care) OR (Maternal-Child Health Centers)) **AND TITLE-ABS-KEY**(dashboard* OR (score card*) OR scorecard* OR (score board*) OR scoreboard* OR graph* OR (visual display*) OR visualisation* OR visualization*) **AND TITLE-ABS-KEY**(quality OR performance OR improv* OR (Quality of Health Care) OR indicator* OR measure*) **AND NOT TITLE-ABS-KEY** graphene	**2,326** *Update: 1,732*
Scopus *Search 2*	**TITLE-ABS-KEY(**Neonatal OR Newborn OR new-born OR Neonate* OR Infant* OR Premature OR Preterm OR Birth OR Childbirth OR Delivery OR Labor OR Labour OR Natal OR Postnatal OR Perinatal OR Prenatal OR Partum OR Postpartum OR Peripartum OR Intrapartum OR Prepartum OR Antepartum OR (Premature Birth) OR (Prenatal Care) OR (Maternal-Child Health Centers)) **TITLE(**dashboard* OR (score card*) OR scorecard* OR (score board*) OR scoreboard* OR graph* OR (visual display*) OR visualisation* OR visualization*) **AND TITLE-ABS-KEY**(quality OR performance OR improv* OR (Quality of Health Care) OR indicator* OR measure*) **AND NOT TITLE-ABS-KEY** graphene	**912** *Update: 644*
Web of Science (Core Collection), Editions: All, *Search 1*	(Neonatal OR Newborn OR new-born OR Neonate* OR Infant* OR Premature OR Preterm OR Birth OR Childbirth OR Delivery OR Labor OR Labour OR Natal OR Postnatal OR Perinatal OR Prenatal OR Partum OR Postpartum OR Peripartum OR Intrapartum OR Prepartum OR Antepartum OR (Premature Birth) OR (Prenatal Care) OR (Maternal-Child Health Centers)) **TITLE** AND (dashboard* OR (score card*) OR scorecard* OR (score board*) OR scoreboard* OR graph* OR (visual display*) OR visualisation* OR visualization*) **TOPIC** AND (quality OR performance OR improv* OR (Quality of Health Care) OR indicator* OR measure*) **TOPIC** NOT (graphene) **TOPIC**	**1,993** *Update: 1,105*
Web of Science (Core Collection), Editions: All, *Search 2*	(Neonatal OR Newborn OR new-born OR Neonate* OR Infant* OR Premature OR Preterm OR Birth OR Childbirth OR Delivery OR Labor OR Labour OR Natal OR Postnatal OR Perinatal OR Prenatal OR Partum OR Postpartum OR Peripartum OR Intrapartum OR Prepartum OR Antepartum OR (Premature Birth) OR (Prenatal Care) OR (Maternal-Child Health Centers)) **TOPIC** AND (dashboard* OR (score card*) OR scorecard* OR (score board*) OR scoreboard* OR graph* OR (visual display*) OR visualisation* OR visualization*) **TITLE**	**690** *Update: 482*

	AND (quality OR performance OR improv* OR (Quality of Health Care) OR indicator* OR measure*) **TOPIC** NOT (graphene) **TOPIC**	
Cochrane Library Reviews	**in Title, Abstract, Keyword** (Neonatal OR Newborn OR new-born OR Neonate* OR Infant* OR Premature OR Preterm OR Birth OR Childbirth OR Delivery OR Labor OR Labour OR Natal OR Postnatal OR Perinatal OR Prenatal OR Partum OR Postpartum OR Peripartum OR Intrapartum OR Prepartum OR Antepartum OR (Premature Birth) OR (Prenatal Care) OR (Maternal-Child Health Centers)) **AND (**dashboard* OR (score card*) OR scorecard* OR (score board*) OR scoreboard* OR graph* OR (visual display*) OR visualisation* OR visualization*) **AND** (quality OR performance OR improv* OR (Quality of Health Care) OR indicator* OR measure*)	**50 Reviews** *Update: 4 Reviews of which none would be eligible*
Total 2010 – 2020	**Identified and included number of articles for the scoping review of this thesis**	**9,073**

Appendix 4.2. Inclusion and Exclusion Criteria

Inclusion criteria	**Exclusion criteria**
Study setting	
Facility level dashboards * *Extended search to Global/national/sub-county/state level), and newborn initiatives or programmes, for example NGOs or research*	- Community level dashboards - Patients or Students as the only users of dashboard (except mixed dashboard audience, e.g. patients and clinicians) - Paper-based reports except they were handed out in addition to electronic visual tools
Low-, middle-, and high-income countries	
Population	
Dashboards related to - Neonatal, facility-based care (birth until first 4 weeks of life) - Reproductive Maternal Newborn and Child Health (RMNCH) only if newborn indicators are reported - Maternal health only if newborn indicators are reported	Dashboards related to other populations than neonates (e.g. babies >= 4 weeks)
Type of indicators	
Any neonatal care indicators *Extended search to indicators around birth related newborn events.*	- Maternal, pregnancy related indicators (only) - Paediatric indicators (babies >= 4 weeks)
Outcome	
Aim of dashboard is any kind of performance or quality of newborn care improvement] – ideally but no strict exclusion criteria if only dashboard design study	Patient reported outcomes
Study types	
All study types.	

* Any information management tool that visually tracks, analyses and displays newborn key performance indicators (KPI), metrics and key data points to monitor the newborn health or specific related processes with electronic tools.

Appendix 4.3. *Overview Included Studies*

#	Authors (Publication year)	Title	Country	Setting	Study design/ Methods	Study objectives	Dashboard Category/ Type
1	**Flanning** C. et al (2010), [148]	Dashboard in neonatology	UK	Neonatal Unit, Craigavon Area hospital, Northern Ireland - One hospital of the National Hospital Service (NHS)	Report / Case study on the implementation of a clinical dashboard in a neonatal unit aimed at monitoring performance and improving patient care by displaying real-time data.	- to explore the benefits of using clinical dashboards in neonatology for improving quality of care and staff satisfaction, - to present advantages of using a neonatal dashboard that displays information.	**Dashboard** on newborn care, including indicators on staffing, environment al cleanliness and mandatory staff training
2	**Sprague**,A. E. et al (2012), [149]	Measuring quality in maternal-newborn care: developing a clinical dashboard	Canada	All maternal-newborn hospitals in Ontario as part of the “Better Outcomes Registry & Network” (BORN). All hospitals involved in the network and that offer maternal-newborn care are using the BORN Information System (BIS), which is an online data collection system the initiative created and that contains information on all births occurring in those hospitals.	This study provides a report of a Phase 1 project to develop a clinical dashboard. Methods used were a **survey & systematic literature review:** The maternal-newborn data elements were aligned with a quality domain framework by the authors. They requested feedback through a survey to assess the relevance and feasibility of making changes. Additionally, they analyzed existing data and literature to help establish provincial benchmarks for	- to identify KPIs to measure quality within maternal-newborn systems and provide automated A&F about KPIs and support quality improvement, - to develop a maternal-newborn dashboard to increase awareness about selected KPIs and support hospitals and care providers in quality improvement.	**Dashboard** / Maternal-Newborn Dashboard

					creating a maternal-newborn dashboard.		
3	**Simms**, R.A. et al (2013), [150]	Development of maternity dashboards across a UK health region; current practice,continuing problems	UK	15 consultant-led maternity units in Southwest England, which is a large and diverse health region.	Mixed-methods design, electronic survey followed by semi-structured interviews	- to assess clinical dashboard developments in line with UK national guidance, - to identify issues faced by maternity units across an entire region and develop quality assurance systems.	**Dashboard** / Maternity dashboard
4	**Gaspar**, J. et al (2013), [151]	Maternal and neonatal healthcare information system: Development of an obstetric electronic health record and healthcare indicators dashboard	Brazil	The study takes place in the scenario of a public maternity hospital of tertiary reference (high-risk pregnancy) bound to a public Brazilian university. Time: August - December 2012	Exploratory survey, simultaneous development of an obstetric EHR and a survey study of indicators to measure the quality of obstetric and neonatal care to inform the development of an electronic tool to automatically calculate a new predetermined set of indicators of maternal and neonatal healthhealthcare	- to propose, implement and evaluate an obstetric Electronic Health Record and to develop a healthcare indicator dashboard, - to evaluate the hypothesis that interdisciplinary work between medicine and computer science is essential to develop an effective obstetric and neonatal software system.	**Dashboard** / Maternal-Newborn Dashboard
5	**Bhardwaj**, S. et al (2014), [152]	Elimination of mother-to-child transmission of HIV in South Africa: rapid scale-up using quality improvement	South Africa	All districts and provinces. Facility, district, provincial and national PMTCT of HIV programmes. Highlighting a district case study in Lejweleputswa district	Intervention/ QI case study, where the National Department of Health developed and implemented an intervention using data-driven processes to understand facility-level bottlenecks to optimising PMTCT implementation and to scale up priority PMTCT actions nationally.	- to describe the process used to develop, implement and monitor a data-driven intervention, - to improve facility, district, provincial and national PMTCT related performance.	**Dashboard** / Visual dashboard and data for action reports

6	**Raban**, M.S. et al (2015), [153]	Impact of an educational intervention and clinical performance dashboard on neonatal bloodstream infections	South Africa	Tertiary-level neonatal unit at Groote Schuur Hospital, Cape Town, South Africa, over 31-months	Comparison of two time periods: Period 1 (January 2010 – April 2011) before any intervention. Period 2 (May 2011 – July 2012) after introduction of an educational intervention and use of clinical performance dashboard.	- to determine whether prospective surveillance of bloodstream infections (BSIs), introduction of an educational intervention and the use of a clinical performance dashboard could reduce BSIs and blood culture contamination rates in a neonatal nursery.	**Dashboard** / Clinical performance dashboard on neonatal bloodstream infections
7	**Temko**, A. et al (2015), [154]	Clinical implementation of a neonatal seizure detection algorithm	Republic of Ireland	Cork University Maternity Hospital, the second largest maternity hospital in Europe	Evaluation of three different ways to visualise detector output and a survey (system validation) on usefulness and accuracy of presented methods among clinical personnel	- to present different ways to visualise outputs for a neonatal seizure detection system and analyse the influence on performance in clinical care, - to explore the algorithm-driven audification of neonatal EEG as an alternative to a visual output.	**Decision support tool** / Visualisal neonatal seizure detection system
8	**Dunn**, S. et al (2016), [155]	A mixed methods evaluation of the maternal-newborn dashboard in Ontario: Dashboard attributes, contextual factors, and facilitators and barriers to use: A study protocol	Canada	All maternal-newborn hospitals in Ontario	Study protocol for a mixed-methods design study with the following methods for each study objective: Interrupted time series (ITS) analysis, key informant interviews with a purposeful sample, a provincial survey, and a multivariable generalised linear mixed effects regression analysis of indicators.	- to evaluate the effect of implementing the dashboard across Ontario, - to explore factors that can explain differences in the use of the the dashboard among hospitals and measure factors potentially associated with differential effectiveness of the dashboard - to identify factors that predict differences in hospital performance.	**Dashboard** / Maternal-Newborn Dashboard
9	**Weiss**, D. et al (2017), [156][156]	Effect of a population-level performance dashboard intervention on	Canada	All maternal-newborn hospitals in Ontario	A quasi-experimental study design was used to assess the impacts of interventions in cases where random allocation is not possible.	- to evaluate the effect of implementation of the dashboard on the rates of six clinical performance indicators in maternal-newborn care across	**Dashboard** / Maternal-Newborn Dashboard

		maternal-newborn outcomes: an interrupted time series study			The interrupted time series (ITS) approach, a strong, quasi-experimental study design was employed to evaluate the effects of interventions when random allocation is not possible.	Ontario hospitals. (which relates to objective 1 from the previously published protocol, Dunn et al. (2016)).	
10	**Diaz**, D. et al (2017), [157]	Visual tools for the exploration of growth data in a cohort of kangaroo infants during their first year of life	Colom-bia	Fundación Canguro - KMC programme. Fundación Canguro has been providing care for newborn premature infants in Colombia since 1994.	Visual analytics project that aimed to assess the development of premature infants during their first year of life. The goal was to determine how variables such as the mother's age, education, weight, height, and income impact the growth process.	- to present a process model or framework suitable for evaluating outcomes of tome-based processes considering the environment (before the event), pregnancy (development of the event), birth (the event), and growth (after the event), - to develop and implement a visual analytics tool that implements visual analytics techniques, - to demonstrate the value of the tool presenting two example case studies	**Visual analytics tool** to evaluate the growth process during the first year of life of premature infants
11	**Szymanski**, L. et al (2017), [158]	Implementation of a multicenter shoulder dystocia injury prevention program	USA	Five geographically dispersed academic medical centers and their affiliated hospitals united by a common insurance carrier	Desciption of development and process of implementation of a shoulder dystocia simulation programme which includes a dashboard to track performance regarding simulation training and clinical and financial outcomes	- to describe the development and implementation process of a shoulder dystocia simulation programme across five medical centers and affiliated hospitals linked to an insurance provider - to promote safe patient, outcomes and reduce preventable injuries, - to reduce medical malpractice claims as a major contributor to medical cost as well as physician satisfaction.	**Dashboard** / Simulation training performance dashboard for shoulder dystocia prevention programme

12	**Patrick**, C. et al (2017), [159]	Improving analysis and reporting for Virginia's Division of Consolidated Laboratory Services newborn screening program	USA	Virginia Newborn Screening programme (NBS), which includes 59 hospitals	Case study for specific programme	- to update the hospital record card with an improved visualisation for sample transit time, a new visualisation of percentage of unsatisfactory samples, a listing of all diagnosis originating from hospital's samples, - to help the Division of Consolidated Laboratory Services (DCLS) more effectively communicate to hospitals their dried blood spot testing (DBS) performance metrics, scale up in response to reporting needs, and better utilise their screening data.	**Hospital report card** with visualisation and information on diagnoses for newborn screening (NBS)
13	**Reszel**, J. et al (2018), [160]	Use of a maternal newborn A&F system in Ontario: a collective case study	Canada	All maternal-newborn hospitals in Ontario	Qualitative study. Collective case study using a qualitative descriptive approach	- to identify factors that facilitate and impede hospitals' use of the dashboard to improve clinical practice and outcomes on KPIs.	**Dashboard** / Materbal-Newborn Dashboard
14	**Rezaei-Hachesu**, P.et al (2018), [161]	The design and evaluation of an AMR system for neonatal intensive care units in Iran	Iran	NICUs in hospitals in Northwestern Iran. The AMR surveillance system was implemented as a pilot software programme in three selected hospitals between Jan and Oct 2017	Developmental study, following previous descriptive studies conducted by the same research team that determined the important axes of AMR surveillance systems. Designing and evaluating surveillance system with user experience questionnaire.	- to design an AMR surveillance system for use in NICUs in northwestern Iranian hospitals to cover information gaps and to considers core elements of biosurveillance and systems, including data collection, analysis, and reporting (using dashboards and geographical information systems).	**Dashboard** / NICU AMR surveillance system with three modules: Data registry, dashboard and clinical decision support modules

15	**Bunch**, K. J. et al (2018), [162]	Developing a set of consensus indicators to support maternity service quality improvement: using Core Outcome Set methodology including a Delphi process	UK	English maternity units (but authors state that the techniques used could equally well be applied in other settings)	Delphi process followed by a face-to-face consensus meeting. The study used a "Core Outcome Set" development methodology which consists of 4 stages: 1) A systematic review to identify reported outcomes, 2) qualitative work to identify outcomes considered important by patients, 3) a Delphi process to prioritise outcomes, and finally 4) a face-to-face consensus meeting.	- to develop a core metric set for monitoring QoC in maternal care, - to investigate whether it was possible to adapt *'Core Outcome Set'* development methodology to identify in a timely, transparent and robust manner, QoC measures that could improve patient experience and outcomes.	**Dashboard** / Maternity dashboard
16	**Jankowska**, M. M. et al (2019), [163]	An Online Geographic Data Visualization Tool to Relate Preterm Births to Environmental Factors	USA	Fresno County, California, had the state's highest preterm birth rate at 9.9%, accounting for 3.7% of California's preterm births. 41.4% of these births occurred in Fresno's densely populated south and west-central areas, where over 70% of pregnant women rely on Medi-Cal health insurance due to low income.	Report	- to discuss the collaborative efforts of several organisations in Fresno County and the state of California to create an online data visualisation and exploration tool by using open software, - to describe preterm birth in Fresno County.	**Online platform** / Online data visualisatio and exploration tool
17	**Patel**, M. S. (2019), [164]	Development and Implementation of Maternity	Oman	Nizwa Hospital, Oman	Prospective, descriptive cross-sectional study conducted between January and December 2017, in	- to develop and implement a maternal care dashboards, - to evaluate the potential of visual displays of clinical performance	**Dashboard** / Maternity dashboard

		Dashboard in Regional Hospital for Quality Improvement at Ground Level: A Pilot Study			patients with high-risk pregnancies who had been referred to Nizwa Hospital	and outcomes for improving the quality of health care at the hospital level.	
18	**Begum**, T. et al (2020), [165]	Perceptions and experiences with district health information system software to collect and utilize health data in Bangladesh: a qualitative exploratory study	Bangla-desh	Two districts of Bangladesh: To choose the study districts they considered 10 different data sets related to RMNCAH services coming from 64 administrative districts of Bangladesh for a one year period	Qualitative study. Key informant interviews (n=11), in-depths interviews (n=23) and focus group discussions (n=2)	- to explore the perceptions, challenges and facilitators and other experiences of various users regarding the implemention of DHIS2, also linked to data collection, analysis and use of the tool - to explore acception of DHIS2 for collecting reproductive, maternal, and newborn health data at various levels.	**Dashboard** / DHIS 2 integrated dashboard

Appendix 4.4. Dashboard KPI Selection

#	Indicator selection process or classification for indicator reporting	Selected and with dashboard/ visualisation tool presented neonatal indicators	Visualisation types/ graph design and visual benchmarks	Frequency of reporting
1	Since there were no standards available at the time, the authors described that the clinical dashboard KPIs were selected in a way that were relevant for the unit which allowed the neonatal unit to monitor factors important to their specific context.	Besides newborn KPIs (that are not reported in detail in this publication), they included other factors such as occupancy rates, environmental cleanliness, and uptake of mandatory staff training. These KPIs were designed to track performance against local, regional, and national standards relevant to their neonatal practice.	A table-traffic light system was used: Target=green, Warning=yellow, Alert=red. Simple and grouped bar plots, and pie charts	Real-time reporting.
2	6 KPIs which were identified during a rigorous development process (to ensure credibility of the BORN Ontario MND). Janakiraman and Eker's criteria for a "good indicator" and the Robson Classification for reporting on Caesarean sections was used. A dashboard subcommittee was established to direct this project, and a survey conducted based on Hermann and Palmer's framework for selecting core quality measures for the purpose of scoring each of the proposed indicators on a scale from 1 (strongly disagree) to 5 (strongly agree). In addition, an open-ended question was used to identify other indicators. At the same time "evidence summaries for each of the potential indicators were developed in collaboration with the Knowledge to Action Research Centre at the Ottawa Hospital Research Institute".	The 6 selected KPIs are: 1. Proportion of newborn screening samples that are unsatisfactory for testing. 2. Proportion of episiotomy in spontaneous vaginal births. 3. Proportion of formula supplementation in term infants whose mothers intended to breastfeed. 4. Proportion of repeat caesarean section in low-risk women not in labour at term with no medical or obstetrical complications done prior to 39-week gestation. 5. Proportion of women delivering at term who had GBS screening at 35–37-week gestation. 6. Proportion of women induced with an indication of post-dates who are less than 41-week gestation at delivery.	A table-traffic light system was used: Target=green, Warning=yellow, Alert=red. For setting an evidence-informed benchmark peer-reviewed literature was examined and current clinical practice within Ontario, and clinical expertise of the BORN committee members was used.	Real-time reporting. Aggregation of monthly observations (e.g., using a quarterly estimate rather than a monthly estimate to achieve a more reliable reflection of clinical practice), and data smoothing (i.e., 3-month moving averages) is built into dashboard.
3	The RCOG was followed for using a traffic light system. Statistics are highlighted red when the upper threshold is breached, requiring	Primarily maternity/pregnancy indicators were displayed. QIs related to newborns were:	A “red, amber, green” (RAG) traffic light system to display data in	No specific information

	immediate investigation; amber to alert staff that potential action may be required to avoid entering the red zone, and green for satisfactory performance.	1. Activity indicators: Births, Instrumental delivery, Caesarean section, Multiple births, Normal births, Neonatal readmissions, Place of birth, Homebirths. 2. Workforce indicators: Midwife to birth ration, supervisor to midwife ratio, 1:1 care in labour. 3. Clinical indicators: Neonatal outcomes: Meconium aspiration, Hypoxic encephalopathy, Birth trauma, Neonatal mortality, Low Apgar score, Ill babies on SCBU. 4. Risk management: Shoulder dystocia. Stillbirths, Breastfeeding.	relation to performance thresholds and statistic displays were used. There was a combination of thresholds from the RCOG guidance, national documents and targets determined locally. None of the units used a dial system as advised by the NHS Connecting for Health Clinical Dashboards Programme.	
4	No specific classification or guidance was followed, but authors discussed that it is "necessary to bring together the clinic-obstetrical experience…and scientific and technological production of health information services", referring to the multidisciplinary need. Before setting up indicators to benchmark quality of care, a database with 126 variables was developed, and variables defined from epidemiological data. After data were collected, they were inserted into an Obstetric EHR (which has been develop for this project specifically). Identified indicators were then grouped in modules offering the user a graphic presentation and therefore a perception of quality care (indicator dashboard within EHR. The indicator dashboard was featured with 70 distinct indicators that approached epidemiological, clinical, administrative and qualitative aspects (defined as essential to evaluate quality of care)	The indicator dashboard contains 70 distinct indicators defined as essential to evaluate the quality of provided care. For this paper the focus was on the following indicators: 1. Caesarean section rate, 2. high-risk patients rate, 3. length of stay, 4. Apgar score after 5th minute below 7, 5. Apgar score after the 5th minute below 7 (adjusted), 6. Breastfeeding in the 1st hour of life rate, 7. Breastfeeding in the 1st hour of life rate (adjusted), 8. Maternal discharge status of patient, 9. Neonatal discharge of patient with seven possible conditions: hospital discharge, hospital transfer to another hospital, exit against medical advice, neonate retained, death before admission, death during admission and neonatal death.	Tables, bar charts, pie charts.	Retrospective preview of indicators, and real-time preview of administrative indicators
5	For this case study in Lejweleputswa district there was an improvement across-the-board in the indicators identified. While no specific	Key indicators monitored from 2010 – 2012 were: 1. ANC 1st visit <20 weeks	Key indicators were scored from 2010 to 2012 using a "traffic light" system to track progress towards	Quarterly data for action reports,

	classification or guidance was reported for indicator selection the following key bottlenecks were identified as a basis: A. Low rate of early antenatal registration (<20 weeks) B. Low rate of CD4+ testing for HIV-positive women (to assess eligibility for ARV treatment) C. Low rate of antenatal clients initiated on PMTCT prophylaxis or on ARV treatment based on eligibility (CD4+ counts >350 cells/µl or <350 cells/µl, respectively)	2. ANC client CD4+ test rate 3. ANC initiated on AZT during ANC rate 4. ANC initiated on HAART rate 5. ANC clients ≥32 weeks retesting rate 6. Postnatal care mother visits within 6 days rate 7. Baby PCR test rate at 6 weeks	the 2012 national target. Green indicated the target was met, amber showed progress without reaching the target, and red signified no progress or a decline. For each indicator, specific action points were discussed and outlined.	including dashboards, were completed at district, provincial and national levels.
6	No specific guideline mentioned. The main focus was on blood culture data and bloodstream infections. Positive blood cultures were classified as contaminants if they detected alpha-haemolytic streptococci, Micrococcus spp., Propionibacterium spp., Corynebacterium spp., Bacillus spp., and sometimes enterococci or Acinetobacter spp., depending on the patient's clinical status. Coagulase-negative Staphylococcus was considered a contaminant if C-reactive protein levels remained <10 mg/L over 24 hours, repeat cultures were negative, or if the clinician did not initiate antibiotics due to the absence of sepsis symptoms in the infant.	Blood culture data/ contamination rates from the microbiology laboratory were used	The tool uses a colour-coded traffic light system, trends, and line charts. Descriptive results are presented as numbers and percentages (%), with a p-value <0.05 considered significant. The colour-coding system uses: green for positive blood cultures below the median, blue for those equalling the median, and red for those exceeding the median. A baseline median was calculated, and this "rolling" median serves as the performance benchmark.	The dashboard serves as a visual aid to demonstrate monthly positive blood cultures and contamination rates and is highlighted and referred to weekly at the neonatal unit staff meeting.
7	Authors followed recommendations for Electroencephalography (EEG) measurement by the International Federation of Clinical Neurophysiology regarding minimum seizure duration window lengths. The system was validated using leave-one-patient-out (LOO) cross-validation, a method recognised for providing an unbiased performance	EEG data were obtained from eight bipolar channel pairs, specifically F4–C4, C4–O2, F3–C3, C3–O1, T4–C4, C4–Cz, Cz–C3, and C3–T3. The data were downsampled to 32 Hz, with an anti-aliasing filter applied at 12.8 Hz. The EEG recordings were segmented into 8-second epochs, with a 50% overlap between consecutive epochs. An 8-second window was selected because hypoxic-ischaemic encephalopathy	EEG plots, aEEG channels. Amplitude, binary and probabilistic line graphs and spatio-temporal colormaps were used. Binary has a threshold benchmark while the others did not.	Real-time

	assessment. This validation utilised a large clinical dataset, including extended multi-channel EEG recordings from 18 neonates with seizures and 20 without. Additionally, the system was tested on a separate dataset of 41 neonates and a larger randomised dataset of 51 full-term neonates.	(HIE) is the most frequent cause of seizures in full-term neonates, and the background EEG in these cases is consistently abnormal. This window length helps prevent the misclassification of brief seizure-like activities, such as short intermittent rhythmic discharges, as seizure events.	A few metrics are commonly used to quantify performance of the neonatal seizure detector: Metrics based on patients, events and epochs.	
8	Authors refer to BORN publication by Sprague et al (#2). In addition, Glantz's recommendation were referenced "to accept lower intervention rates if higher rates do not improve outcomes was a guiding principle, as well as other risk assessment techniques such as restricting patient populations to more homogeneous subgroups."	Referred to #2	Referred to #2	Referred to #2
9	Small numerators - high variability in monthly rates of newborn screening samples, therefore for this indicator quarterly time interval was used	Referred to # 2, 8	Referred to # 2, 8	Referred to # 2, 8
10	Pre-collected data, though the authors categorized them into four categories based on the process model	47 measured variables ranging from environment (mother's age, weight, etc.), pregnancy (number of hospitalisations, diet, etc.), birth (birth weight, head circumference, etc.), growth (anthropometrics, etc.)	Various visual codes were considered, including position, colour, and form. For instance, "mother's age" was displayed as a bar graph (mark: line, channel: vertical position), a simple and commonly used format for easy understanding. More complex visualisations were also created, such as a graph showing the evolution of feeding types during a child's first year, incorporating different marks (lines, circles) and channels (position, size, colour) in one graph. Personalised time series were used as benchmarks,	Retrospective analysis, no frequency involved

			incorporating growth time units, and WHO standards.	
11	An observation checklist was used by programme facilitators that assesses the simulation technical and behaviour skills along with teamwork and communication. Authors report that overall metrics architecture will be continuously improved, and the metrics will allow the comparison of individual institutions to other MCIC institutions. A future goal of the programme is to improve shared learning through MCIC institutional and peer programme information exchanges.	Dashboard performance measures include: 1. Total percent of labour and delivery staff successfully completing simulation training once every 2 years (attending physicians, resident physicians, nursing staff, other) 2. Total % of shoulder dystocia cases per total vaginal deliveries 3. Total % shoulder dystocia cases with signs of BPI immediately after delivery 4. Total % of shoulder dystocia cases with signs of BPI at discharge 5. Total % of shoulder dystocia cases with signs of BPI 12 months post discharge 6. Claims rate/ number related to BPI 7. Claims cost related to BPI	No specific information	The simulation training programme is to complete at least once every 2 years.
12	Pre-collected data. Not very sure - the false positive section has degrees of freedom information	Infant (patient ID, birth date, birth time, ethnicity, weight in grams, transfused, feed type, feed other, zip code), sample (sample ID, collection date, received date), submitter (submitter ID, submitter name), test (MSUD, PKU etc 36 in all), diagnosis (diagnosis, diagnosis date, sample ID)	Line, GIS (generated in R using ggplot2) Benchmark: DCLS targets, its previous metrics, and the performance of other hospitals, as well as false positive degree of freedom	Monthly
13	*No specific information but referencing # 2, 8, 9.*	See # *2, 8, 9.*	See # *2, 8, 9.*	See # *2, 8, 9.*
14	Study is based on BI requirement analysis and a predefined comprehensive antimicrobial surveillance framework compatible with WHO global surveillance strategy (GLASS). A checklist was developed to extract BI requirements. These requirements were embedded in the software modules. Scrum for Research (SCORE) methodology was used to follow agile development of the surveillance system, and the Scrum framework was	7 major AMR surveillance topics in eight managerial axes are considered: 1. surveillance methods in AMR, 2. priority specimens, 3. priority samples, 4. priority pathogens 5. microbial agents, 6. test methods and reporting scheduling,	Tables, bar charts, pie charts, GIS. To compare the system with those used in similar previous research, the research team employed a UEQ benchmark dataset consisting of 9905 records from 246 studies on software systems.	Real-time, weekly, monthly depending on Scrum product owner (PO).

	customized based on the limitations of the research team. Nine iterative and incremental steps were defined in the final methodology.	7. recommended datasets and reporting protocols/tools. The main entities for the dashboard are: a complete view of clinical and laboratory data elements (Final Flat), antimicrobial agents (Antimicrobial Agent), GIS data elements (IRN_adm2), resistance statistical data (Combination TBL), and selected pathogens to monitor (National pathogens); Two types of results are displayed: features associated with neonatal mortality and resistance patterns.		
15	A) Metrics were identified through a review of existing dashboards in England, covering six domains: antenatal, maternal, neonatal, mental health, public health, and workforce. These were assessed in the first phase of the Delphi process. B) Experts from maternity care were organised into three panels—service design, service provision, and public. Each received an information pack and a link to the data collection website for the Delphi process. C) In Phase 1, participants scored candidate metrics (1-9) for importance in monitoring maternal care quality and could suggest additional metrics, which were reviewed by two assessors. D) Phase 2 of the Delphi process. In total 144 metrics were included and finally 125 distinct metrics across six domains were identified and were listed. In addition, 19 new metrics were nominated and added at phase 2. After scoring these, 79 metrics met the criteria for consideration and for discussion at	The 14 metrics included for the final dashboard set are: 1. smoking rate at booking 2. rate of birth without intervention 3. caesarean section delivery rate in Robson group 1 women 4. caesarean section delivery rate in Robson group 2 women 5. caesarean section delivery rate in Robson group 5 women 6. third- and fourth-degree tear rate among women delivering vaginally 7. rate of postpartum haemorrhage of ≥1500 ml 8. rate of successful vaginal birth after a single previous caesarean section 9. smoking rate at delivery 10. proportion of babies born at term with an Apgar score <7 at 5 minutes 11. proportion of babies born at term admitted to the neonatal intensive care unit 12. proportion of babies readmitted to hospital at <30 days of age 13. breastfeeding initiation rate 14. breastfeeding rate at 6–8 weeks	No specific information, graphical and numerical presentation	Monthly monitoring

	a consensus meeting. Finally, 14 metrics were selected for the dashboard.			
16	The tool integrates data on births, environmental exposures, sociodemographic factors, the environment, and healthcare access. The topic-driven front end provides a guided experience, for users to explore predefined topics (e.g., demographics, environmental pollution) with curated data, limited to 15 variables per screen to ensure ease of use for non-experts. According to the authors, the data-driven back end enables users to analyse preterm birth variables by region and examine their relationships with other variables without restrictions on quantity or type.	The birth cohort file this tool is based on and reports on includes comprehensive details on maternal and infant characteristics, obtained from linked hospital discharge records, birth certificates, and infant death records. Additional variables derived from the linked birth certificate and hospital data encompass race/ethnicity, maternal age, parity, country of birth, pre-pregnancy weight, height, insurance status, smoking habits, diabetes, hypertension, anaemia, maternal education, reported drug use, diagnosed infections, and mental illness. The file also records the number of previous caesarean sections and the interpregnancy interval.	The tool opens with a video introduction, guiding users through 8 key topical areas designed to facilitate exploration of preterm birth and geographically organised data. The platform offers 4 main visualisation functionalities, for users to explore data by geographic region, by specific indicators, by relationships between indicators, and through correlation matrices. It also allows multiple ways to examine data relationships, using dynamic visuals such as pie charts, histograms, ranked associations with other indicators, and heat-map matrix displays.	Unclear whether real-time feedback or just one-off visualisation
17	The selection of quality indicators was based on the prototype of clinical outcomes from the Royal College of Obstetricians and Gynaecologists (RCOG). The guidance by the NHS Institute for Innovation and Improvement [21] for what a good indicator is was used and measures were aligned with “The Good Practice Series” on the Maternity Dashboard. In addition to pre-selected measures, local parameters were adapted based on clinical observations The dashboard was placed in the delivery suite so that frontline staff could review the display monthly.	Selected parameters which are broadly categorised were: **For clinical activities:** 1. Deliveries 2. Admission to the antenatal ward (excluding direct admission 3. to the delivery suite) 4. Outpatient department appointment 5. Instrumental deliveries 6. Caesarean section rate **For maternal measures:** 1. Induction of labour 2. Workforce 3. Midwife/patient ratio 4. Supervisor/midwife ratio	Traffic-light colour-coded system, however, looks more like a scorecard.	Monthly

		5. Eclampsia 6. Intensive care unit admission 7. Severe postpartum haemorrhage 8. Third-degree perineal tear 9. Shoulder dystocia 10. Hematomas 11. Postpartum hysterectomy 12. Others (e.g., near miss/mortality) **For neonatal outcomes:** 1. Low five-minute Apgar score (< 7) 2. Perinatal asphyxia 3. Meconium aspiration syndrome 4. Stillbirth 5. Stillbirth with diabetes 6. Early neonatal death.		
18	Since it is a national information system used on different levels of the health system, DHIS2 collects a myriad of data. In addition to RMNCAH indicators, the system captures information on human resources, medicine, and logistics data. For this study, the authors refer to "RMNCAH indicators" retrieved by the DHIS2.	According to the authors, users recommended that the DHIS2 platform needs to be set up to automatically generate data for key RMNCAH indicators. A pop-up box with each indicator's definition, calculation, and disaggregation options should be added to provide quick support and standardise data collection. They also suggest translating the software into Bangla to improve understanding of instructions and RMNCAH indicators.	No specific information. Data can be visualised individually byDHIS2 users.	Ideally, real-time, but this has been acknowledged as a challenge. Aggregated data are being reported monthly and feedback meetings based on reports monthly.

Appendix 4.5. Additional Research Articles

Overview of additional research articles from 2021 – 2024, of which #19 shows to present additional evidence linked to the BORN dashboard, that was included in the review through earlier publications

#	Year	Country	Dashboard Category	Dashboard Type	Dashboard Aim	Authors	Title
18	2021	Australia	Dashboard	Maternal-Newborn Dashboard	To monitor perinatal outcomes during the COVID-19 pandemic	Hui L, et al. [254]	Collaborative maternity and newborn dashboard (CoMaND) for the COVID-19 pandemic: a protocol for timely, adaptive monitoring of perinatal outcomes in Melbourne, Australia
19	2021	Canada	Dashboard	Maternal-Newborn Dashboard	To provide a clinical audit & feedback system for QI	Murphy M, et al. [255]	Data Resource Profile: Better Outcomes Registry & Network (BORN) Ontario
20	2023	Australia	Dashboard	Newborn dashboard	To monitor and analyse neonatal physiological signals at Neonatal Intensive Care Unit (NICU)	Lauw C, et al. [231]	Development of an Interactive Dashboard to Analyse Physiological Signals in the Neonatal Intensive Care Unit

Chapter V Quantitative Analysis Appendices

Appendix 5.1. Survey Questions JISC

Assessing feedback on performance for neonatal care improvements

0%

0% complete

Welcome

What is this survey about?
This survey is part of a mixed-methods study that aims to learn how to design and implement hospital performance dashboards and other feedback tools for improving quality of care for newborns in Low- and Middle-Income Countries (LMICs). We hope to gather feedback from end-users like yourselves to tailor the design of dashboards and other feedback on performance tools so that we meet your needs. This is NOT a test of your ability. Instead, we want to learn from you how useful different tool features are. Your contribution will help to make feedback on performance in hospitals better and more useful for you and others.
Your participation in this survey is voluntary and takes about 20 minutes or more depending on speed. All responses will be anonymous, and they will be used exclusively for academic purposes. If you want to learn more about the ethics of this study and how your data is protected and processed, please find more details here.

This survey is confidential, and you will be reimbursed for data use, if you fully complete the survey. PLEASE SHARE THIS SURVEY AMONGST YOUR NETWORK!

Contact
If you have any comments or questions, or you wish to be involved more, please contact christiane.hagel@ndm.ox.ac.uk. Thank you in advance for taking the time!

By starting this survey you give your consent to participate in this survey and to processing of the information as outlined above/ in the detailed information sheet. To start the survey, please click NEXT.

Jisc Online surveys

Dashboard Survey (Data collection 1)

Response ID	Start date	Completion date
[redacted]	5 May 2022, 07:28 (BST)	5 May 2022, 11:58 (BST)

1	What is your age?	31 - 40 years old
2	What is your gender?	[redacted]
3	Please let us know in which country you currently work in.	Kenya
4	Do you work in a position that is directly related to neonatal care?	Yes
5	What term decribes best your current main role?	Clinical-hospital role
6	What is your role in the non-hospital context?	
6.a	If you selected Other, please specify:	
7	...how often do you use a computer?	
8	...how often do you use a smartphone on which you can read emails and documents?	
9	...how often do you use WhatsApp or other messaging services to give/ receive feedback on your/ your team's work performance?	
10	What is your role at the hospital?	Consultant
10.a	If you selected Other, please specify:	
11	...how often do you use a computer?	Every day
12	...how often do you use a smartphone on which you can read emails and documents?	Every day

13	...how often do you use WhatsApp or other messaging services to give/ receive feedback on your/ your team's work performance?	Every day

14	Which one describes your role best?	
14.a	If you selected Other, please specify:	

15	...how often do you use a computer?	

16	...how often do you use a smartphone on which you can read emails and documents?	

17	...how often do you use WhatsApp or other messaging services to give/ receive feedback on your/ your team's work performance?	

18	Which of these tools or approaches are you most involved in using when it comes to giving/receiving hospital's performance based on neonatal healthcare indicators? Select only 1 that you use most for giving/receiving feedback and respond to the related questions below. If none of these exist where you work, please indicate and continue with question 19.	My role does not involve performance indicators based on healthcare
18.a	If you selected Other, please specify:	
18.b	Please rate the selected one above in terms of its usefulness for improving quality of care. Not relevant if you don't use any of the above.	
18.b.1	Usefulness for quality improvement	
18.b.1.a	Rate from 1 (not useful) - 7 (very useful)	7
18.c	Please rate how you feel about the selected one above in terms of whether you would need more training to make use of this tool for quality of care improvement. Not relevant if you don't use any of the above.	
18.c.1	Training how to use the tool	
18.c.1.a	Rate from 1 (I don't need more) - 7 (I need much more)	5
18.c.2	Training how to interpret data, graphs and/or tables	
18.c.2.a	Rate from 1 (I don't need more) - 7 (I need much more)	1
18.c.3	Training how to use feedback to change practice to improve quality of care	
18.c.3.a	Rate from 1 (I don't need more) - 7 (I need much more)	3

19	If you could select one, which of these tools or approaches would you prefer over the other for giving/receiving feedback on performance? Select only 1 that you'd prefer even though it might not exist where you work.	My role does not involve performance indicators based on healthcare
19.a	If you selected Other, please specify:	

20	In general, what frequency of performance feedback based on quality of care indicators do you think is useful for your....	
20.1	...individual performance?	Weekly
20.2	...team's performance?	Bi-weekly
20.3	...hospital's/ organisation's performance?	Monthly

21	The following neonatal outcome indicators are measured and presented. Please rate for each indicator how important it is for you to receive/give feedback on the same for understanding quality of newborn hospital care. Or indicate if you think this question is irrelevant for you/ your work.	
21.1	Facility births (number)	
21.1.a	Rate from 1 (not important at all) - 7 (very important), or indicate if irrelevant for you/ your work.	7
21.2	Mortality, number of newborn deaths by cause	
21.2.a	Rate from 1 (not important at all) - 7 (very important), or indicate if irrelevant for you/ your work.	7
21.3	Stillbirth rate	
21.3.a	Rate from 1 (not important at all) - 7 (very important), or indicate if irrelevant for you/ your work.	4
21.4	Pre-discharge mortality rate	
21.4.a	Rate from 1 (not important at all) - 7 (very important), or indicate if irrelevant for you/ your work.	5
21.5	Newborns admitted for inpatient care (number)	
21.5.a	Rate from 1 (not important at all) - 7 (very important), or indicate if irrelevant for you/ your work.	6

22	The following process indicators are measured and presented. Please rate for each indicator how important it is for you to receive/give feedback on the same for understanding quality of newborn hospital care. Or indicate if you think this question is irrelevant for you/ your work.	
22.1	Birthweight always documented	
22.1.a	Rate from 1 (not important at all) - 7 (very important), or indicate if irrelevant for you/ your work.	6
22.2	Early initiation of breastfeeding performed	

22.2.a	Rate from 1 (not important at all) - 7 (very important), or indicate if irrelevant for you/ your work.	6
22.3	Success of early initiation of Kangaroo Mother Care (KMC) for premature babies	
22.3.a	Rate from 1 (not important at all) - 7 (very important), or indicate if irrelevant for you/ your work.	6
22.4	Success of continuous positive airway pressure (CPAP) therapy	
22.4.a	Rate from 1 (not important at all) - 7 (very important), or indicate if irrelevant for you/ your work.	6
22.5	Use of chlorhexidine	
22.5.a	Rate from 1 (not important at all) - 7 (very important), or indicate if irrelevant for you/ your work.	5
22.6	Use of vitamin K	
22.6.a	Rate from 1 (not important at all) - 7 (very important), or indicate if irrelevant for you/ your work.	6

23	If you feel like different information in form of key performance indicators would be more useful, please let us know which one this could be:	

24	Who would you say are the main users of the newborn information above? Please name up to three.	Hospital Managers
24.a	If you selected Other, please specify:	

25	Please rate for the highlighted main key performance issue of hospital A whether it improved over time or not.	
25.1	Hospital A's recent performance vs. past performance	
25.1.a	Improved: Rate from 1 (not at all) - 7 (very much)	3
25.2	Hospital A's recent performance vs. other hospital's recent average performance	
25.2.a	Improved: Rate from 1 (not at all) - 7 (very much)	1
25.3	Other hospital's recent average performance vs. hospital A's past performance	
25.3.a	Improved: Rate from 1 (not at all) - 7 (very much)	2

26	How suitable do you think are the following items of the table above for providing/receiving feedback on performance: Please note: Benchmarking means the process of measuring key metrics and comparing them to others. The figure legend refers to the accompanied table to explain the table/colours/interpretation.	
26.1	Overall design of the table	
26.1.a	Suitable: Rate from 1 (not at all) - 7 (very much)	6

26.2	Benchmarking	
26.2.a	Suitable: Rate from 1 (not at all) - 7 (very much)	5
26.3	Terminology	
26.3.a	Suitable: Rate from 1 (not at all) - 7 (very much)	5
26.4	Colour scheme	
26.4.a	Suitable: Rate from 1 (not at all) - 7 (very much)	6
26.5	Figure legend	
26.5.a	Suitable: Rate from 1 (not at all) - 7 (very much)	6

27	Please rate the entire table above in terms of (a) how easy it is to interpret information, and (b) how useful feedback in form of this design/table is for quality of care improvements:	
27.1	(a) Easy to interpret	
27.1.a	Rate from 1 (not at all) - 7 (very much)	5
27.2	(b) Useful	
27.2.a	Rate from 1 (not at all) - 7 (very much)	5

28	Please have a look at hospital A's admissions of newborns and answer the question below. Note: Referral into hospital newborn units (NBU) means those referred from other facilities. Born before arrival (BBA) means home deliveries or not born in any healthy facility. In how many months were referrals greater than BBA?	
28.1	No. of months:	
28.1.a		1

29	Please rate the entire visualisation above in terms of (a) how easy it is to interpret information, and (b) how useful feedback in form of this graph is for quality of care improvements:	
29.1	(a) Easy to interpret	
29.1.a	Rate from 1 (not at all) - 7 (very much)	5
29.2	(b) Useful	
29.2.a	Rate from 1 (not at all) - 7 (very much)	6

30	Familiarise yourself with this graph that presents hospital A's newborn mortality trends by birth weight and answer the question below. Please rate how easy it is to interpret from this graph whether mortality is improving, staying the same or getting worse based on the information presented:	
30.1	Mortality birth weight category 1 - <1.5kg	
30.1.a	Easy to interpret: Rate from 1 (not at all) - 7 (very easy)	5

30.2	Mortality birth weight category 2 - <2.5 kg from Jan - April	
30.2.a	Easy to interpret: Rate from 1 (not at all) - 7 (very easy)	4
30.3	Mortality in total	
30.3.a	Easy to interpret: Rate from 1 (not at all) - 7 (very easy)	6

31	Please rate the entire visualisation above in terms of (a) how easy it is to interpret information, and (b) how useful feedback in form of this graph is for quality of care improvements:	
31.1	(a) Easy to interpret	
31.1.a	Rate from 1 (not at all) - 7 (very much)	6
31.2	(b) Useful	
31.2.a	Rate from 1 (not at all) - 7 (very much)	6

32	Familiarise yourself with this graph presenting the documentation of signs and symptoms of severe illness and answer the question below. If you had to rate performance over time based on this graph, how would you rate quality of care improvement for the following	
32.1	Temperature from May - July	
32.1.a	Performance Scale: Rate from -3 to -1 (declined a lot - declined), 0 (stayed the same), 1 to 3 (improved - improved a lot). Or indicate if difficult to tell.	2
32.2	Heart rate from Sept - Oct	
32.2.a	Performance Scale: Rate from -3 to -1 (declined a lot - declined), 0 (stayed the same), 1 to 3 (improved - improved a lot). Or indicate if difficult to tell.	Difficult to tell
32.3	SPO2 from May - August	
32.3.a	Performance Scale: Rate from -3 to -1 (declined a lot - declined), 0 (stayed the same), 1 to 3 (improved - improved a lot). Or indicate if difficult to tell.	Difficult to tell
32.4	In total	
32.4.a	Performance Scale: Rate from -3 to -1 (declined a lot - declined), 0 (stayed the same), 1 to 3 (improved - improved a lot). Or indicate if difficult to tell.	Difficult to tell

33	Please rate the entire visualisation above in terms of (a) how easy it is to interpret information and (b) how useful feedback in form of this graph is for quality of care improvements	
33.1	(a) Easy to interpret	
33.1.a	Rate from 1 (not at all) - 7 (very much)	3
33.2	(b) Useful	
33.2.a	Rate from 1 (not at all) - 7 (very much)	5

34	What is the neonatal mortality rate in August? Please enter number without percent sign (%).	15

35	What is the lowest admission between Sept. '20 and Sept. '21?	119

36	What is the mortality rate in the normal birth weight category?	9%

37	Please rate the entire visualisation above in terms of (a) how easy it is to interpret information and (b) how useful feedback in form of this graph is for quality of care improvements	
37.1	(a) Easy to interpret	
37.1.a	Rate from 1 (not at all) - 7 (very much)	4
37.2	(b) Useful	
37.2.a	Rate from 1 (not at all) - 7 (very much)	4

38	According to the graph above, what are the two most common primary diagnoses at discharge/death? Please choose two from the options below.	• Infections • Prematurity

39	By what % were other primary diagnoses documented at discharge/death?	None of the above

40	What is the number of babies that died? Please select from the options below.	I don't know

41	What number of babies receiving phototherapy died? Please select from the options below.	I don't know

42	How many admitted babies survived? Please select from the options below.	Around 254

43	How many admitted babies survived? Please enter a number or text.	Don't know

44	Please rate the entire visualisation above in terms of (a) how easy it is to interpret information and (b) how useful feedback in form of this graph is for quality of care improvements	
44.1	(a) Easy to interpret	
44.1.a	Rate from 1 (not at all) - 7 (very much)	1
44.2	(b) Useful	

44.2.a	Rate from 1 (not at all) - 7 (very much)	5

45	Which device was not available? Please select only one from the list below.	Catheter

46	How many babies were taken care of during the night?	48

47	Please rate the entire visualisation above in terms of (a) how easy it is to interpret information and (b) how useful feedback in form of this graph is for quality of care improvements	
47.1	(a) Easy to interpret	
47.1.a	Rate from 1 (not at all) - 7 (very much)	2
47.2	(b) Useful	
47.2.a	Rate from 1 (not at all) - 7 (very much)	1

48	Have you ever used a neonatal routine report of the Clinical Information Network (CIN) for your work? Please find an example of how this report may look like here.	Yes, I have used it.
48.a	If YES, do you feel it is useful?	Yes
48.b	Provide further feedback here if you would like to do so (e.g. on the complexity, length, frequency, accuracy, whether it provides a fair picture of things, fairness, etc. or anything else that you might want to share):	

49	Have you ever used the NEST360 dashboard for neonatal care for your work? Please find an example of how this dashboard may look like here.	No, I have never used it.
49.a	If YES, do you feel it is useful?	No
49.b	Provide further feedback here if you would like to do so (e.g. on the complexity, length, frequency, accuracy, whether it provides a fair picture of things, fairness, etc. or anything else that you might want to share):	

50	Does your work involve either patient care or patient quality improvement?	Yes

51	We appreciate your feedback to understand how confident you feel in general within your role related to quality of care improvement. Please complete and rate the following sentence: I feel confident...	
51.1	(a) ...to assess performance information on quality of care	
51.1.a	Rate from 1 (not at all confident) - 7 (very confident)	6

51.2	(b) ...how to use performance information to improve quality of care	
51.2.a	Rate from 1 (not at all confident) - 7 (very confident)	6
51.3	(c) ...to have the skills to improve quality of care	
51.3.a	Rate from 1 (not at all confident) - 7 (very confident)	5
51.4	(d) ...that I am able to achieve improvements in my hospital/ ward	
51.4.a	Rate from 1 (not at all confident) - 7 (very confident)	3
51.5	(e) ...to know what feedback on performance and quality of care means	
51.5.a	Rate from 1 (not at all confident) - 7 (very confident)	4

52	I want to be more involved in this study for interviews and agree to be contacted by the investigators of this study:	No

53	Kindly provide your phone number here for reimbursement:	

Appendix 5.2. Complete Anonymised CIN Report 2021

NEONATAL ROUTINE REPORT

Of a general teaching and referral Hospital

July '21 - September '21

1. BACKGROUND INFORMATION

- This report is part of a series of quarterly feedback reports given to each hospital in the country.
- The Ministry of Health have made efforts to improve medical record keeping in clinical areas by providing and promoting the use of standardised paediatric and newborn admission record forms in keeping with WHO recommendations.
- Data are collected from medical records (Neonatal Admission Record-NAR) of all sick patients admitted in the New Born Unit (age and spectrum of diseases depend on the hospitals' organization) at the time the patient is discharged to examine documentation of history, physical examination, diagnosis, laboratory investigations, treatment and discharge plans.
- Data are collected in an electronic tool - REDCap - by a trained data clerk under close supervision by the hospital medical records department
- Data quality assurance is done by an external research team and de-identified data are shared with this team to enable the production of these feedback reports
- Data are presented for the most recent 3 months for each hospital and preceding 3 months. Also provided are data on the hospital's cumulative performance
- Data also show changes in performance over time allowing hospitals to monitor their trends and examine the effect of efforts at improvement.
- Color key:

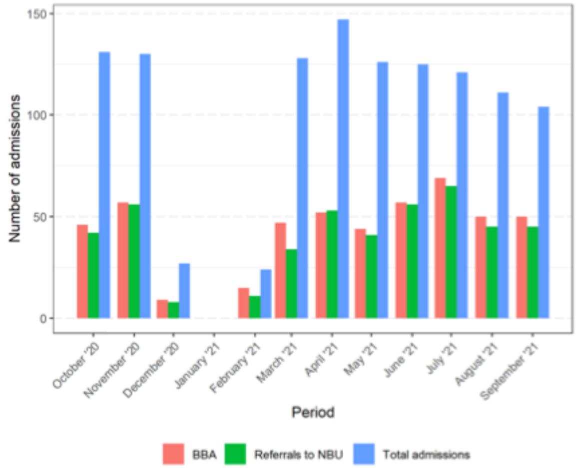
Color	Intepretation	Range
	Excellent Performance	>90%
	Good Performance	80 - 90%
	Some Performance	60 - 79%
	Poor Performance	<60%

1. SUMMARY

TABLE 1: KEY ISSUES

		General Teaching and Referral Hospital	General Teaching and Referral Hospital	All Other Hospitals
		(Jul '21 - Sep '21)	(Mar '21 - Jun '21)	(Jul '21 - Sep '21)
Proportion documented	Birth Weight	362/375 (96.53%)	476/493 (96.55%)	6786/6929 (97.94%)
	Admission Weight	351/375 (93.6%)	464/493 (94.12%)	4396/6929 (63.44%)
	Discharge Weight	274/375 (73.07%)	368/493 (74.65%)	5518/6929 (79.64%)
In-patient Monitoring For Birth weight less than (<) 2.5 kg and Length of stay >=7 days	Weight monitoring (>=1 documented)	99/99 (100%)	90/90 (100%)	1249/1249 (100%)
	Weight monitoring (>=2 documented)	99/99 (100%)	90/90 (100%)	1190/1249 (95.28%)
	Weight monitoring (>=3 documented)	99/99 (100%)	90/90 (100%)	1190/1249 (95.28%)

FIGURE 1: ADMISSIONS



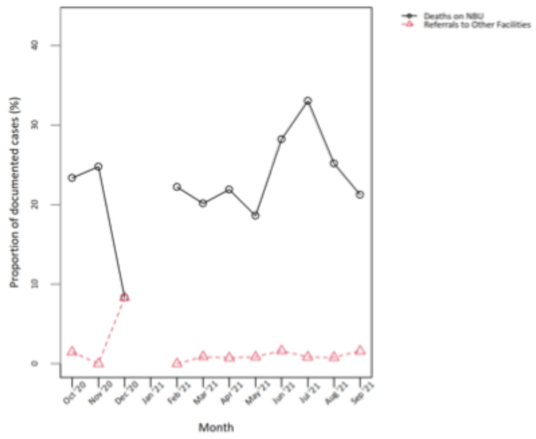
Note:

These are admissions during the month that had been discharged by the end of the reporting period.

Referral into NBU – those referred from other facilities.

BBA – Home deliveries or not born in any healthy facility.

FIGURE 2: PROPORTION OF MORTALITY AND REFERRAL ON TOTAL DISCHARGES



Note:

This section on outcomes for those whom it is documented.

Deaths in NBU - Those who died.

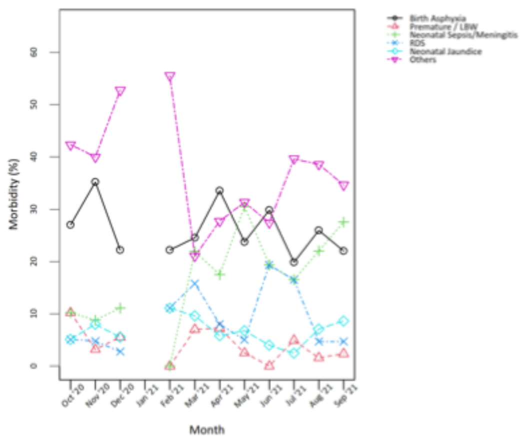
Referred to Other Facilities- Those who were referred to Other Facilities/Hospitals

2. PATIENT CHARACTERISTICS

TABLE 2: PATIENT CHARACTERISTICS AT ADMISSION

		General Teaching and Referral Hospital	General Teaching and Referral Hospital	All Other Hospitals
		(Jul '21 - Sep '21)	(Mar '21 - Jun '21)	(Jul '21 - Sep '21)
Total cases		n = 375	n = 493	n = 6929
Proportion of NAR used	NAR used	374/375 (99.73%)	493/493 (100%)	6654/6929 (96.03%)
Birth weight	Documented cases	362/375 (96.53%)	476/493 (96.55%)	6786/6929 (97.94%)
	<1 kg	13/362 (3.59%)	16/476 (3.36%)	175/6786 (2.58%)
	1 - <1.5 kg	52/362 (14.36%)	47/476 (9.87%)	490/6786 (7.22%)
	1.5 - <2 kg	72/362 (19.89%)	67/476 (14.08%)	960/6786 (14.15%)
	2 - <2.5 kg	41/362 (11.33%)	71/476 (14.92%)	891/6786 (13.13%)
	2.5 - 4 kg	172/362 (47.51%)	259/476 (54.41%)	3915/6786 (57.69%)
	>4 kg	12/362 (3.31%)	16/476 (3.36%)	355/6786 (5.23%)
Age at Admission	Information on age available	375/375 (100%)	493/493 (100%)	6898/6929 (99.55%)
	Admitted on day of birth	294/375 (78.4%)	390/493 (79.11%)	5845/6898 (84.73%)
	Day after birth to 7 days	346/375 (92.27%)	454/493 (92.09%)	6742/6898 (97.74%)
	>7 days	28/375 (7.47%)	39/493 (7.91%)	94/6898 (1.36%)
Discharge Weight	Documented cases	274/375 (73.07%)	368/493 (74.65%)	5518/6929 (79.64%)
Length of stay (days)	Documented cases	374/375 (99.73%)	493/493 (100%)	6921/6929 (99.88%)
	Mean	10.66	8.24	7.63
	Median(Max - Min)	6(0-75)	5(0-91)	4(0-89)
Gestation at birth	Documented cases	350/375 (93.33%)	473/493 (95.94%)	6061/6929 (87.47%)
	<28 wks	19/350 (5.43%)	23/473 (4.86%)	194/6061 (3.2%)
	28 - <34 wks	104/350 (29.71%)	90/473 (19.03%)	1003/6061 (16.55%)
	34 - <37 wks	57/350 (16.29%)	86/473 (18.18%)	964/6061 (15.9%)
	37 - 42 wks	166/350 (47.43%)	268/473 (56.66%)	3850/6061 (63.52%)
	> 42 wks	4/350 (1.14%)	6/473 (1.27%)	50/6061 (0.82%)
Gender	Documented cases	375/375 (100%)	493/493 (100%)	6902/6929 (99.61%)
	Male	207/375 (55.2%)	287/493 (58.22%)	3925/6902 (56.87%)
	Female	168/375 (44.8%)	206/493 (41.78%)	2974/6902 (43.09%)
Maternal HIV status	Documented cases	275/375 (73.33%)	427/493 (86.61%)	6246/6929 (90.14%)
	Positive	11/275 (4%)	18/427 (4.22%)	347/6246 (5.56%)
	Negative	264/275 (96%)	409/427 (95.78%)	5899/6246 (94.44%)
Mode of delivery	Documented cases	375/375 (100%)	493/493 (100%)	6862/6929 (99.03%)
	Spontaneous vaginal (SVD)	284/375 (75.73%)	330/493 (66.94%)	4188/6862 (61.03%)
	Caesarean section (C/S)	82/375 (21.87%)	157/493 (31.85%)	2510/6862 (36.58%)
	Other	9/375 (2.4%)	6/493 (1.22%)	164/6862 (2.39%)
Discharge Weight	Documented for LOS>=7days OR birth weight <2kg	133/184 (72.28%)	152/206 (73.79%)	2085/2542 (82.02%)
Outcome	Documented cases	375/375 (100%)	493/493 (100%)	6927/6929 (99.97%)
	Alive	272/375 (72.53%)	378/493 (76.67%)	5903/6927 (85.22%)
	Dead	99/375 (26.4%)	110/493 (22.31%)	901/6927 (13.01%)
	Others*	4/375 (1.07%)	5/493 (1.01%)	123/6927 (1.78%)

FIGURE 3: DISEASE PATTERN TRENDS AT DISCHARGE

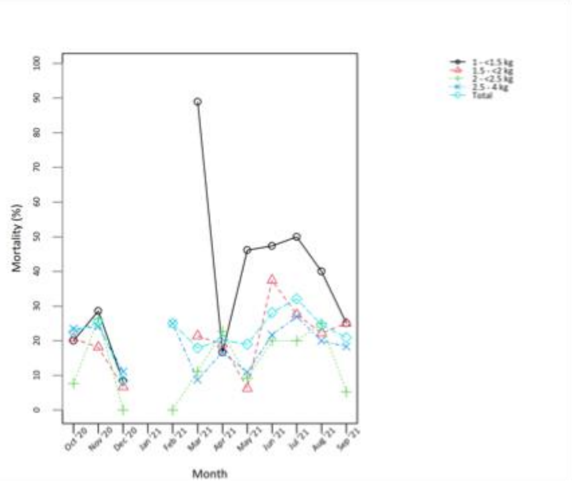


Note:

This section shows trends of discharge diagnosis as documented by clinicians at discharge. These are monthly disease episodes so that a patient with more than one diagnosis will contribute to the count for each diagnosis.

3. MORTALITY

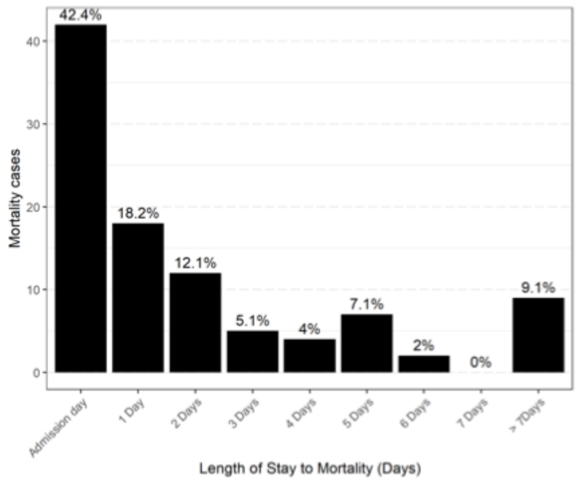
FIGURE 4: MORTALITY BY BIRTH WEIGHT



Note:

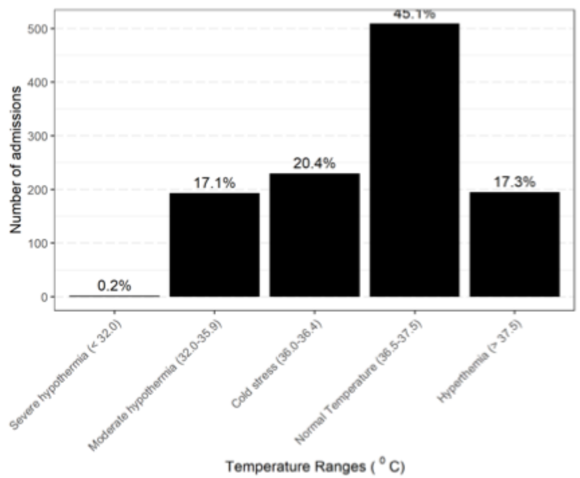
Gap in the trend indicates there were no cases for that specific period in the group hence no outcomes.

FIGURE 5: PROPORTIONS OF LENGTH OF STAY IN MORTALITY CASES



This is derived from the current reporting period data (Oct '20 to Sep '21)

FIGURE 8: TEMPERATURE RANGES



This section shows information on temperature as documented during admission during the reporting period (Oct '20 to Sep '21).

4. ANTIBIOTIC USE

TABLE 3: ANTIBIOTIC USE

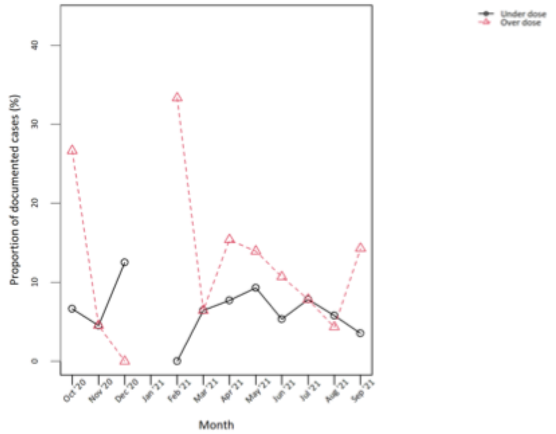
		General Teaching and Referral Hospital	General Teaching and Referral Hospital	All Other Hospitals
		(Jul '21 - Sep '21)	(Mar '21 - Jun '21)	(Jul '21 - Sep '21)
Antibiotic prescriptions	Total cases	116/375 (30.93%)	109/493 (22.11%)	3506/6929 (50.6%)
	Penicillin prescription only	0/116 (0%)	6/109 (5.5%)	59/3506 (1.68%)
	Gentamicin prescription only	0/116 (0%)	0/109 (0%)	14/3506 (0.4%)
	Penicillin & Gentamicin prescription	116/116 (100%)	102/109 (93.58%)	3380/3506 (96.41%)
	Other Antibiotics*	0/116 (0%)	1/109 (0.92%)	53/3506 (1.51%)

Note:

The table gives prescription information on made the date of admission only.

Other Antibiotics*: this includes Amikacin, Ampicillin, Ceftriaxone and Ceftazidime.

FIGURE 9: PENICILLIN DOSAGES



Note:

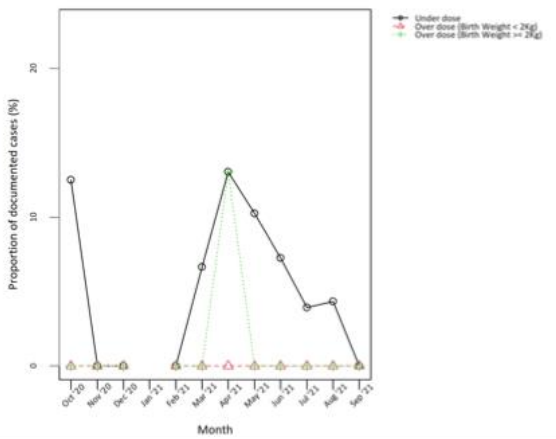
Dosage was calculated using the formula Dosage = dose prescribed/weight

Under dose - a prescription of less than 40,000iu/kg

Overdose - a prescription >60,000 iu/kg

Correct dose - a prescription >=40000 & <=60000 iu/kg

FIGURE 10: GENTAMICIN DOSAGES



Note:

Dosage was calculated using the formula Dosage= dose prescribed/weight.

Underdose - a prescription <2.4mg/kg for weight below 2kg or <4mg/kg for weight >=2kg.

Overdose - a prescription >3.6mg/kg for weight below 2kg or <6mg/kg for weight >=2kg.

5. ENTERAL FEEDING

TABLE 4: ENTERAL FEEDS PRESCRIPTION AT ADMISSION

		General Teaching and Referral Hospital	General Teaching and Referral Hospital	All Other Hospitals
		(Jul '21 - Sep '21)	(Mar '21 - Jun '21)	(Jul '21 - Sep '21)
Type of feed	Total cases on first day of life	294/375 (78.4%)	390/493 (79.11%)	5845/6929 (84.36%)
	Proportion of children receiving both iv and feeds	0/294 (0%)	4/390 (1.03%)	338/5845 (5.78%)
	Total cases without feed/fluid prescribed	57/294 (19.39%)	127/390 (32.56%)	1567/5845 (26.81%)
Type of feed	Total cases with feed prescription	0/294 (0%)	1/390 (0.26%)	725/5845 (12.4%)
	Expressed Breast Milk (EBM) Only	0/0 (0%)	1/1 (100%)	415/725 (57.24%)
	Formula Only	0/0 (0%)	0/1 (0%)	177/725 (24.41%)
	Mixed feeding(EBM + Formula)	0/0 (0%)	0/1 (0%)	113/725 (15.59%)
	Others*	0/0 (0%)	0/1 (0%)	21/725 (2.9%)
Information on age and weight available*	Cases with volume documented*	0/0 (0%)	1/1 (100%)	635/725 (87.59%)
Volume in 24 hrs	Correct dose*	0/0 (0%)	0/1 (0%)	514/635 (80.94%)
	Under dose*	0/0 (0%)	1/1 (100%)	51/635 (8.03%)
	Over dose*	0/0 (0%)	0/1 (0%)	70/635 (11.02%)

Note:

The table gives information on enteral feeds prescribed at admission for children admitted on the day of birth and prescribed enteral feeds without intravenous fluids.

Others* - those prescribed others feed except EBM and formula.

7. NURSING MONITORING

TABLE 6: NURSES MONITORING VITAL SIGNS

		Kakamega County General Teaching and Referral Hospital	Kakamega County General Teaching and Referral Hospital	All Other Hospitals
		(Jul '21 - Sep '21)	(Mar '21 - Jun '21)	(Jul '21 - Sep '21)
Vital signs chart present		342/375 (91.2%)	339/493 (68.76%)	5654/6929 (81.6%)
Nurses Charting of Vital signs	Pulse rate	336/342 (98.25%)	311/339 (91.74%)	5117/5654 (90.5%)
	Temperature	337/342 (98.54%)	332/339 (97.94%)	4972/5654 (87.94%)
	Pulse oximetry	330/342 (96.49%)	268/339 (79.06%)	5074/5654 (89.74%)
	Respiratory rate	335/342 (97.95%)	306/339 (90.27%)	4625/5654 (81.8%)
	Feed monitoring	0/0 (0%)	1/1 (100%)	373/725 (51.45%)
	Fluid monitoring	144/177 (81.36%)	125/187 (66.84%)	1934/2650 (72.98%)

Note:

This section gives information on monitoring done by nurses at any point during the hospital stay and documented in the various monitoring charts (vital signs chart, weight chart and fluid monitoring chart).

6. INTRAVENOUS FLUID USE

TABLE 5: INTRAVENOUS FLUID PRESCRIPTION AT ADMISSION

		General Teaching and Referral Hospital	General Teaching and Referral Hospital	All Other Hospitals
		(Jul '21 - Sep '21)	(Mar '21 - Jun '21)	(Jul '21 - Sep '21)
Type of fluid	Total cases on first day of life	294/375 (78.4%)	390/493 (79.11%)	5845/6929 (84.36%)
	Total cases with IV fluid prescription*	177/294 (60.2%)	187/390 (47.95%)	2650/5845 (45.34%)
	10% D/W	177/177 (100%)	187/187 (100%)	2628/2650 (99.17%)
	5% D/W	0/177 (0%)	0/187 (0%)	7/2650 (0.26%)
	Ringers Lactate	0/177 (0%)	0/187 (0%)	3/2650 (0.11%)
	N/ Saline	1/177 (0.56%)	4/187 (2.14%)	24/2650 (0.91%)
Information on age and weight available*	Cases with volume documented*	173/177 (97.74%)	181/187 (96.79%)	2339/2650 (88.26%)
Volume in 24 hrs	Correct dose*	155/173 (89.6%)	158/181 (87.29%)	2045/2339 (87.43%)
	Under dose*	5/173 (2.89%)	8/181 (4.42%)	113/2339 (4.83%)
	Over dose*	13/173 (7.51%)	15/181 (8.29%)	181/2339 (7.74%)

Note:

First day of life with fluid prescription only (no feeds prescribed).

The table gives information on intravenous fluids prescribed at admission for children admitted on the day of birth and prescribed intravenous fluids without enteral feeds.

Dose calculation was done in cases where the time of fluid administration was 24hours and total volume prescribed was not greater than 500mls. Where iv fluid dosage= (total volume of iv fluids prescribed/birth weight).

Cases with volume documented* : Where volume and frequency of administration is documented.

Correct dose*: fluids dosage >=**64** & <= **96** ml/kg/day for birth weight <1.5kg or fluids dosage of >=**48** & <=**72** ml/kg/day for a birthweight >=1.5 kg

TABLE 7: SUPPORTIVE CARE

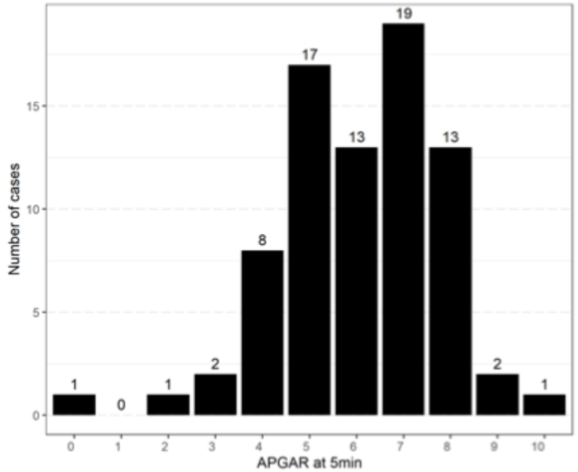
	Kakamega County General Teaching and Referral Hospital	Kakamega County General Teaching and Referral Hospital	All Other Hospitals
	(Jul '21 - Sep '21)	(Mar '21 - Jun '21)	(Jul '21 - Sep '21)
APGAR documentation	375/375 (100%)	493/493 (100%)	6929/6929 (100%)
CPAP	7/375 (1.87%)	18/493 (3.65%)	466/6929 (6.73%)
CPAP in RDS* Cases	4/87 (4.6%)	11/96 (11.46%)	365/2017 (18.1%)
KMC* in babies with birth weight <2kg surviving to discharge	56/89 (62.92%)	62/79 (78.48%)	840/1152 (72.92%)

Note:

RDS*: A clinical diagnosis of respiratory distress syndrome at admission or discharge

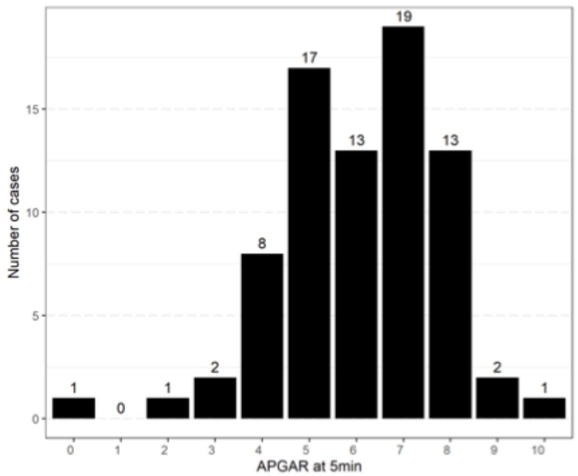
KMC*: Proportion of children prescribed Kangaroo Mother Care among all children discharged alive with a birthweight of less than 2kg.

FIGURE 11: APGAR AT 5 MINUTES FOR BIRTH ASPHYXIA CASES



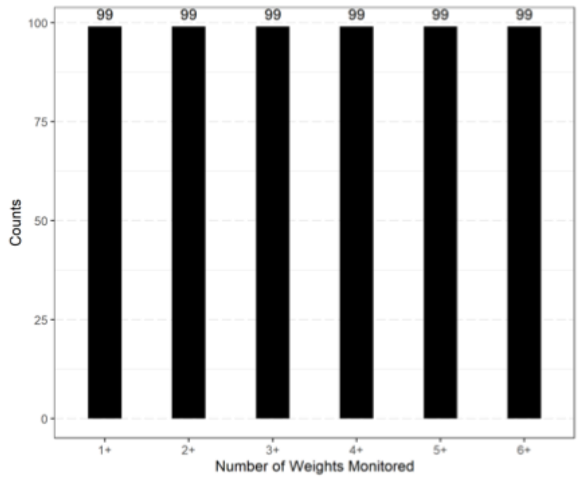
This section gives information of APGAR score for children with a clinical diagnosis of birth asphyxia at admission or discharge.

FIGURE 11: APGAR AT 5 MINUTES FOR BIRTH ASPHYXIA CASES



This section gives information of APGAR score for children with a clinical diagnosis of birth asphyxia at admission or discharge.

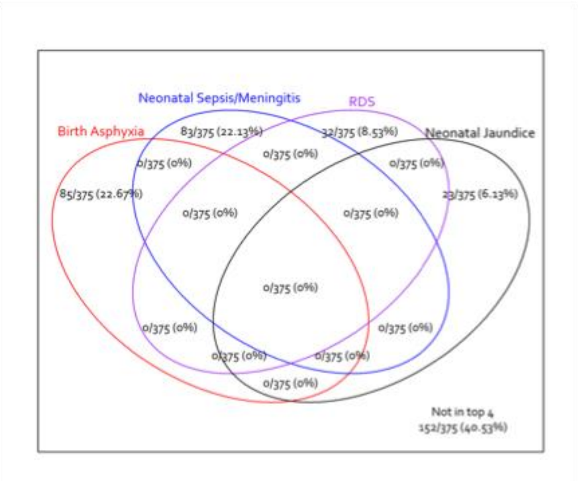
FIGURE 12: POST ADMISSION WEIGHT MONITORING



This section gives information of post admission weight capture. Birth weight, admission weights and post-discharge weights are excluded.

8. APPENDICES

FIGURE A.1: VENN DIAGRAM OF DISEASE PATTERNS AT DISCHARGE

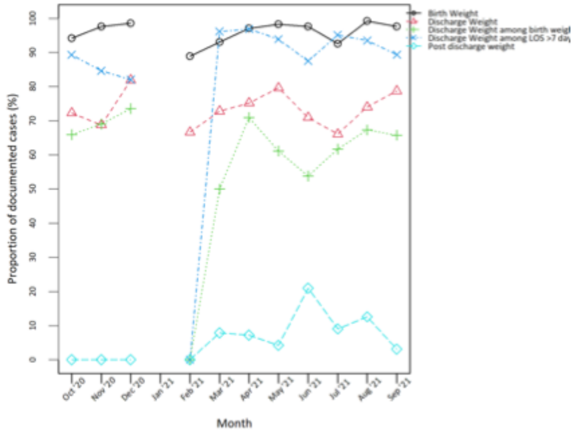


Note:

This is derived from the current reporting period data (Oct '20 to Sep '21)

These are disease episodes i.e., the number of documented diagnoses so that for instance a patient with low birth weight and respiratory distress syndrome would contribute a count to each diagnosis.

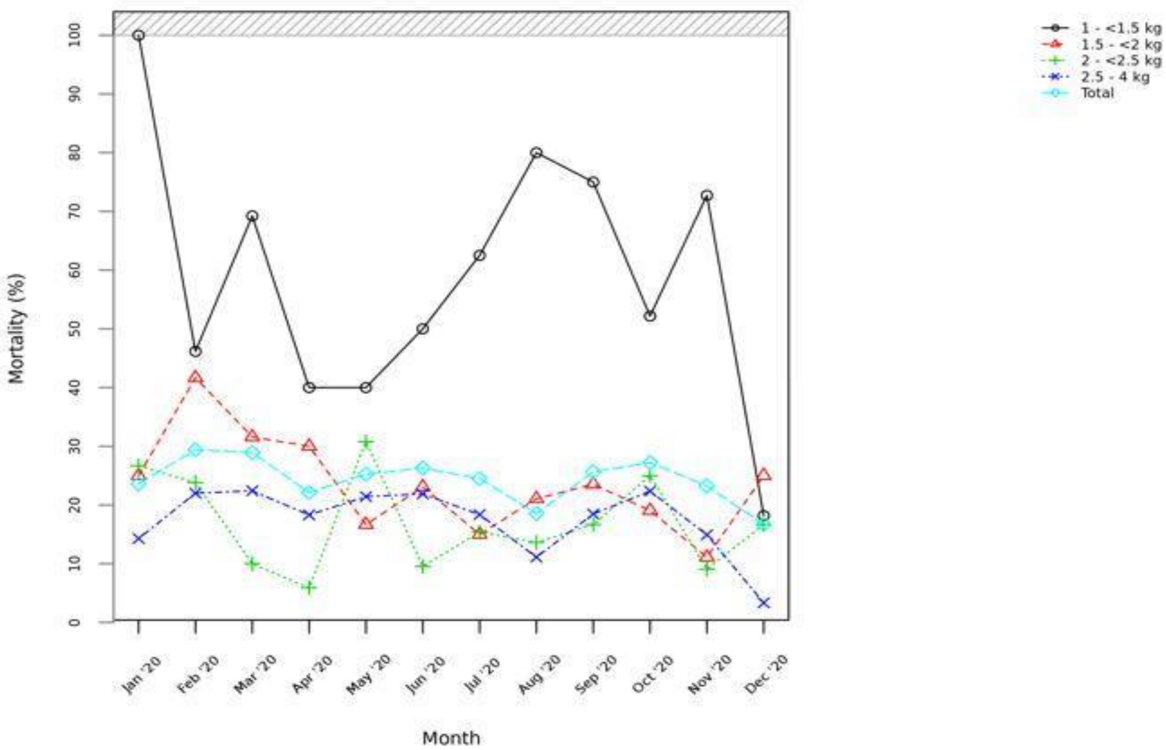
FIGURE A.2: DOCUMENTATION TRENDS FOR WEIGHTS



Appendix 5.3. Excluded CIN Graph

Survey CIN Line Chart Figure on newborn mortality trends by birth weight for survey questions 30 – 31.

This line graph was part of the survey but was excluded from the analysis after the survey questions were retrospectively found to be unsuitable for providing information on the interpretation success rate. However, a similar line chart using performance measures is included, see Figure 8.



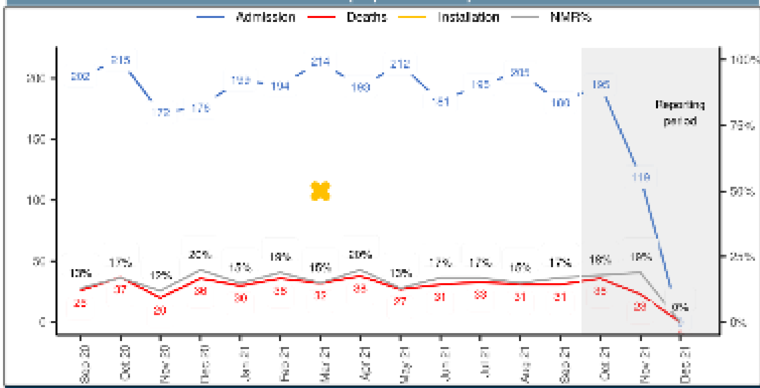
Appendix 5.4. NEST Dashboard

Quality Improvement Dashboard
For Small and Sick New born Care

Facility Name/ID: 240 Reporting Period Oct - Dec 2021 Country: Kenya

Admissions & Mortality

Trends (Sep '20-Dec '21)

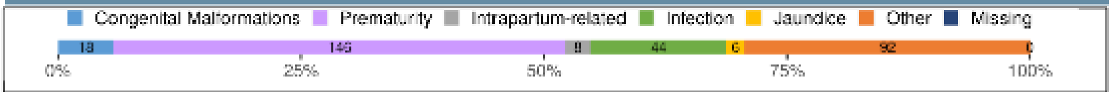


Birthweight Specific (Oct - Dec 2021)

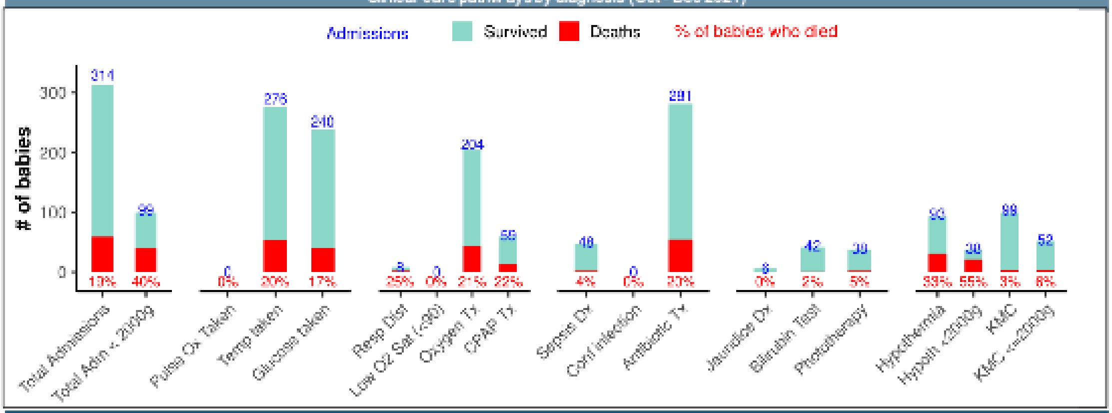
Birthweight	Number of Admissions	Mortality Rate
<1000	20	95%
1001-1500	32	44%
1501-2000	58	12%
2001-2500	47	13%
2501-4000	138	9%
4001+	18	6%
Not recorded	1	0%
Total	314	19%
Median length of stay		5 days
% inborn		81%

Quality of Clinical Care

Primary diagnosis at discharge/death by ICD principles (Oct - Dec 2021) n = 314



Clinical care pathways by diagnosis (Oct - Dec 2021)



Health System Input as of Last QI or HFA Visit (Oct-12-2021)

Human Resources

Personnel	Day of Visit	Night Before Visit
Clinical Officers	0	0
Nurses	5	2
Doctors	2	1
Biomeds assigned	Not recorded	Not recorded
Day baby:nurse	48.5	
Night baby:nurse	48.2	

Infrastructure (space, power, etc.)

Bed capacity (normal)	65
Bed occupancy rate	73.84615326
Power cuts >30min	Not recorded

Device Functionality & Consumable Availability

Oxygen: Concentrator	Oxygen: Tank	Warmers: Radiant	Warmers: Incubator	Digital weighing scale
Functional: 5	Functional: 4	Functional: 8	Functional: 12	Functional: Not recorded
Available: 5	Available: 4	Available: 8	Available: 12	Available: Not recorded
Nasal prongs (neonate): Available		Probes: Available	Probes: Not available	
Nasal prongs (infant): Available				

Digital Thermometer
Functional: 2
Available: 2

Glucometer	CPAP	Suction Pump	Pulse Oximeter	Phototherapy
Functional: 2	Functional: 5	Functional: 5	Functional: 3	Functional: 7
Available: 2	Available: 5	Available: 5	Available: 3	Available: 8
Glucometer strips: Available	CPAP nasal prongs: Available	Catheters: Not available	Probes: Available	

Lightmeter
Functional: 1
Available: 1

Totals may be affected by missing data.

Appendix 5.5. Survey Handout for Tested Data Visuals

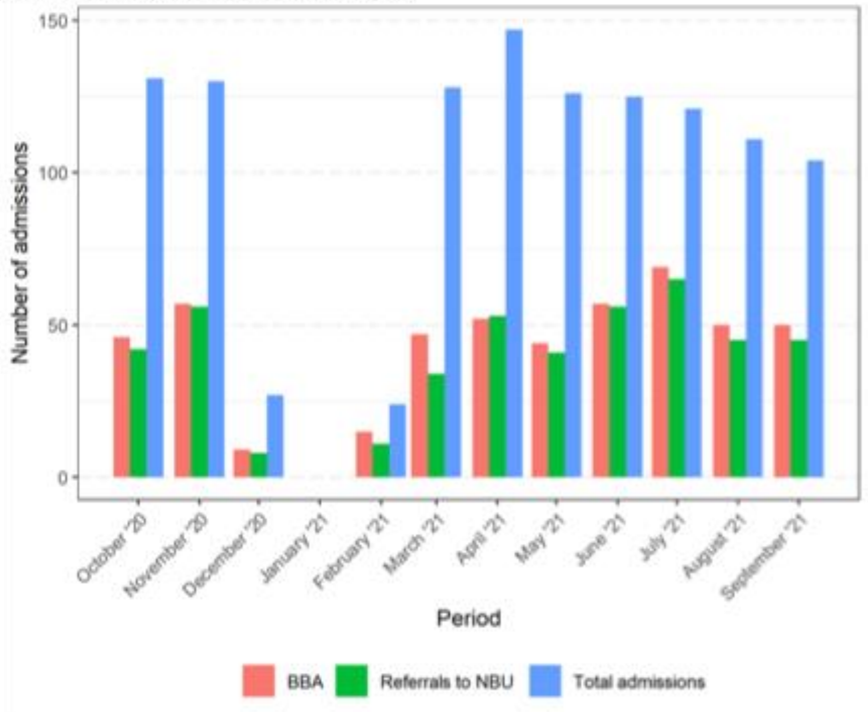
Survey on feedback tools for quality improvement of newborn hospital care

Handout supporting your responses

If you have any comments or questions, or you wish to be involved more, please contact christiane.hagel@ndm.ox.ac.uk. Thank you in advance for taking the time!

2. Figure on admissions of newborns for survey questions 28 - 29

How to design feedback visualisations (2 of 4)



Note: Referral into hospital newborn units (NBU) means those referred from other facilities. Born before arrival (BBA) means home deliveries or not born in any healthy facility. In how many months were referrals greater than BBA?

1. Table on key performance issues of hospital A, for survey question 25-27

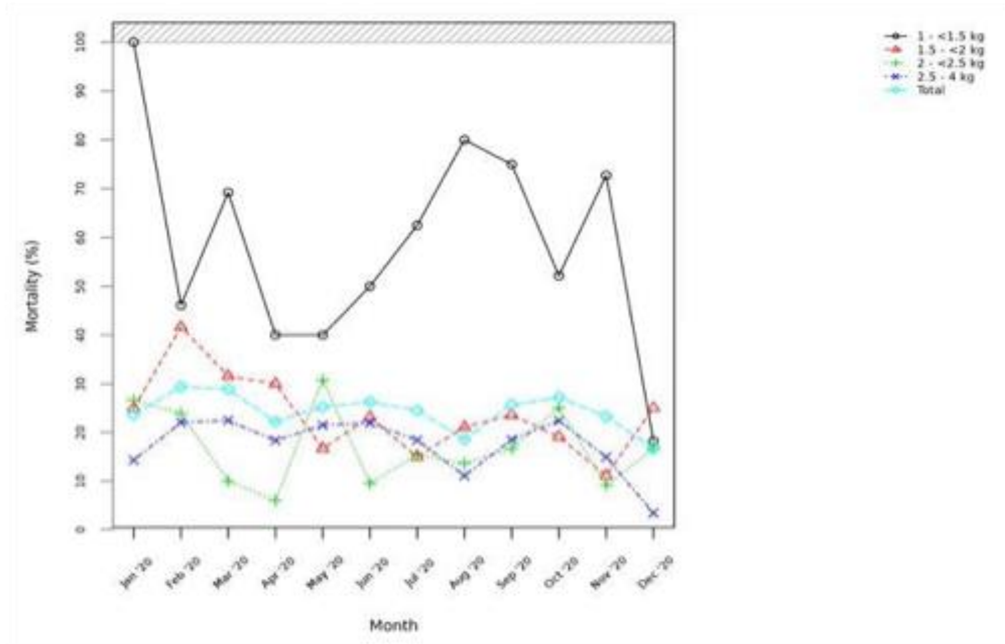
How to design feedback visualisations (1 of 4)

		Hospital A	Hospital A	All Other Hospitals
		Mar '21 - Jun '21	Jul '21 - Sep '21	Jul '21 - Sep '21
Total cases		n = 493	n = 375	n = 6929
Proportion documented	Birth Weight	476/493 (96.55%)	362/375 (96.53%)	6786/6929 (97.94%)
	Admission Weight	464/493 (94.12%)	351/375 (93.6%)	4396/6929 (63.44%)
	Discharge Weight	368/493 (74.65%)	274/375 (73.07%)	5518/6929 (79.64%)
In-patient Monitoring for Birth weight less than (<) 2.5 kg and Length of stay >=7 days	Weight monitoring (>=1 documented)	90/90 (100%)	99/99 (100%)	1249/1249 (100%)
	Weight monitoring (>=2 documented)	90/90 (100%)	99/99 (100%)	1190/1249 (95.28%)
	Weight monitoring (>=3 documented)	90/90 (100%)	99/99 (100%)	1190/1249 (95.28%)
Proportion of Newborn admission records (NAR) used	NAR used	493/493 (100%)	374/375 (99.73%)	6654/6929 (96.03%)

Colour	Interpretation	Range
	Excellent Performance	>90%
	Good Performance	80 - 90%
	Some Performance	60 - 79%
	Poor Performance	<60%

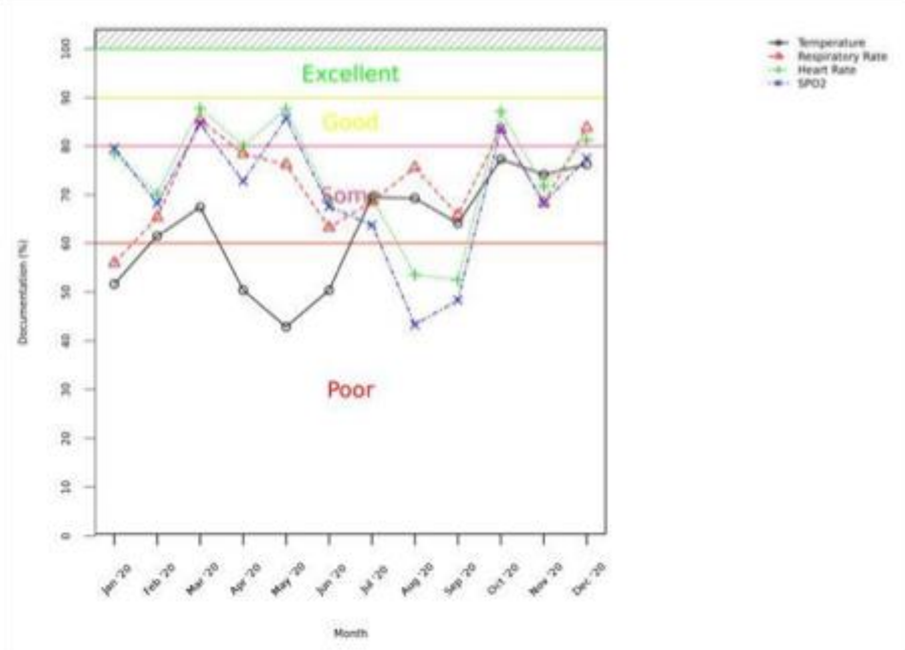
3. Figure on newborn mortality trends by birth weight for survey questions 30 - 31

How to design feedback visualisations (3 of 4)



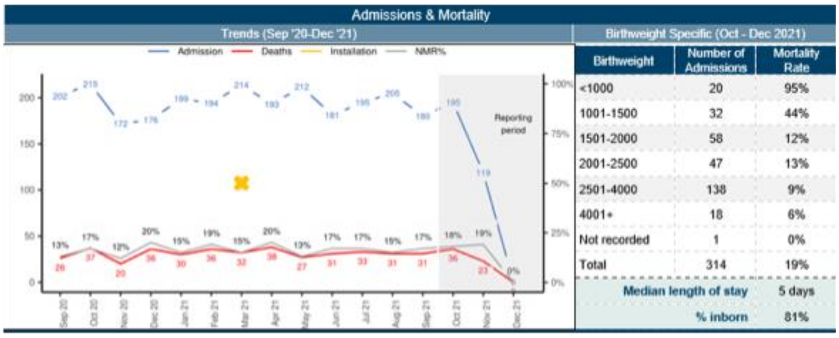
4. Figure on documentation of signs and symptoms of severe illness for survey question 32 - 33

How to design feedback visualisations (4 of 4)



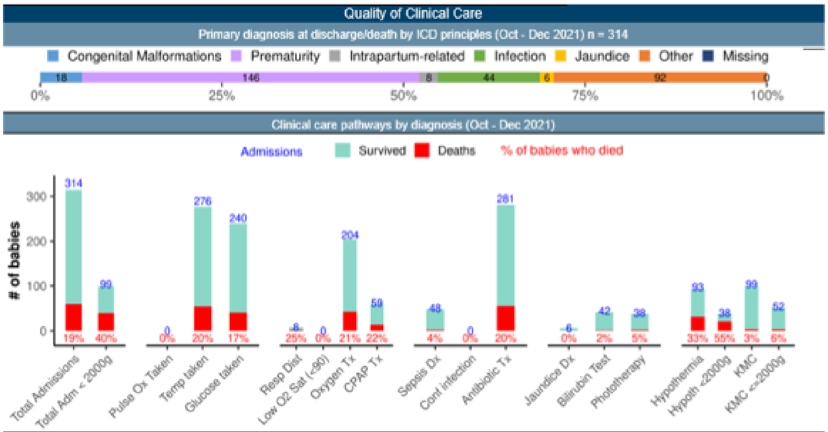
5. Figure on admissions and mortality trends for survey questions 34 - 37

Other feedback designs (1 of 3)



6. Figure on quality of clinical care for survey questions 38 – 43

Other feedback designs (2 of 3)



7. Figure on health system inputs for survey questions 44 - 46

Other feedback designs (3 of 3)

Health System Input as of Last QI or HFA Visit (Oct-12-2021)							
Human Resources			Device Functionality & Consumable Availability				
Personnel	Day of Visit	Night Before Visit	Oxygen		Warmers		Digital weighing scale
Clinical Officers	0	0	Concentrator	Tank	Radiant	Incubator	Functional: Not recorded
Nurses	5	2					Available: Not recorded
Doctors	2	1	Functional: 5	Functional: 4	Functional: 8	Functional: 12	Digital Thermometer
Biomeds assigned	Not recorded	Not recorded	Available: 5	Available: 4	Available: 8	Available: 12	
Day baby:nurse	48:5		Nasal prongs (neonate): Available		Probes: Available	Probes: Not available	Functional: 2
Night baby:nurse	48:2		Nasal prongs (infant): Available				Available: 2
Infrastructure (space, power, etc.)			Glucometer	CPAP	Suction Pump	Pulse Oximeter	Phototherapy
Bed capacity (normal)	65		Functional: 2	Functional: 5	Functional: 5	Functional: 3	Functional: 7
Bed occupancy rate	73.84615326		Available: 2	Available: 5	Available: 5	Available: 3	Available: 8
Power cuts >30min	Not recorded		Glucometer strips: Available	CPAP nasal prongs: Available	Catheters: Not available	Probes: Available	Lightmeter
							Functional: 1
							Available: 1

Appendix 5.6. Survey Participant Information Sheet

ONLINE SURVEY

Study Title: The use of data to improve quality of neonatal hospital care in LMICs and help strengthen health systems

Dear participant,

We appreciate your interest in participating in this implementation process and technology evaluation study to investigate how feedback with dashboards helps to improve quality of neonatal hospital care in low and middle income countries (LMICs). With your help we would like to evaluate current design, graph and delivery types used to present hospital newborn data to give feedback on performance. The survey takes about 10-15 minutes. You can pause the survey at any time and complete at a later moment. We recommend to complete all questions in one session. No background knowledge is required. There are no foreseeable disadvantages or risks in participating in this research. However, we kindly ask you to participate in this study during a time outside your working shift(s).

Please note that the participation is voluntary. You can withdraw at any point for any reason before submitting your answers by pressing the 'Exit' button/ closing the browser. The information you have provided until this point, will be discarded and will not be used in summarizing the research. Your responses will be anonymised and we will not share your personal information. If you wish to get reimbursed for data use and/or participate more in this study, e.g. by being interviewed, you can share your contact details at the end of the survey. Your information will be kept confidential at any time.

Please take the time to read the **Participant Information Sheet** on next pages. If you have any further questions or problems accessing the survey please send an email to christiane.hagel@ndm.ox.ac.uk

Best regards,

Christiane Hagel

MPH, DPhil Candidate in Clinical Medicine
Nuffield Department of Medicine, University of Oxford, UK

Participant Information Sheet

Study title: The use of data to improve quality of neonatal hospital care in LMICs and help strengthen health systems

This is an implementation process and technology evaluation study is part of a mixed-methods study to investigate how feedback with dashboards helps to improve quality of neonatal hospital care in LMICs

1. Who is carrying out this study?

The principal investigator is Ms. Christiane Hagel who is a doctoral research student at the Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine at the University of Oxford. This project is being completed under the supervision of Prof. Mike English, Prof. Sassy Molyneux and Dr. Chris Paton and in collaboration with other investigators from the KEMRI Wellcome Trust Research Programme's Clinical Information Network (CIN) and the NEST360 project. More information on NEST360 can be found here.

Institutions	Investigators
University of Oxford, Nuffield Department of Medicine, Centre for Tropical Medicine and Global Health, Oxford, United Kingdom	Ms. Christiane Hagel, Dr. Chris Paton
KEMRI-Wellcome Trust Research Programme, Nairobi, Kenya, a government organisation that carries out medical research to find better ways of preventing and treating illness in the future for everybody's benefit.	Ms. Joyline Jepkosgei, Dr. Timothy Tuti, Prof. Sassy Molyneux, and Prof. Mike English

2. Why is this research being conducted?

In this research, we want to learn more about how data is used for improving quality of neonatal hospital care in low and middle income countries (LMICs). In particular, we want to examine how well actors from different health sector target groups understand and interpret different data-driven feedback displays and forms (e.g. presented with dashboards), and what their efforts are to understand feedback. This includes to ask these actors about their needs, examining what feedback they would benefit from, what their preferences are for different indicators or types of feedback, and how it motivates them or helps to know what or how to act on it.

3. Why have I been invited to take part?

We are interested in understanding how you use information on neonatal key performance indicators given to you with different information visualisations and feedback tools/dashboards. We feel that your experience as a person who is engaged with the development or use of feedback systems on neonatal indicators (or neonatal dashboards) can contribute to our understanding and knowledge of its potential improvement.

4. Do I have to take part and what will happen if I refuse to participate?

No. Please note that participation is voluntary. You can ask questions about the research before deciding whether or not to take part. If you do agree to take part, you may withdraw yourself from the study, without giving a reason and without negative consequences, by advising me/us of this decision. Please note that all questions of the survey are optional and you can choose the answer option "prefer not to say". You can change your mind at any time and withdraw at any point for any reason before submitting your answers by pressing the 'Exit' button/ closing the browser. The information you have provided until this point of the survey, will be discarded and will not be used in summarizing the research results. You can also withdraw at a later stage after you submitted all answers. The deadline by which you can withdraw all information you have contributes to this research is 30 April 2023.

5. What will happen if I take part in the research?

If you are happy to take part in the research, you will be asked to give answers to our online questionnaire/ survey. The link to the survey will be sent to you by email. The survey takes about 10 - 15 minutes. You can pause the survey at any time and complete at a later moment. However, we recommend to complete all questions in one session. No background knowledge is required. If you wish to participate more in this study, you can share your contact details at the end of the survey. Your information will be kept confidential at any time.

6. Are there any disadvantages (or risks) and advantages to me of taking part?

There are no foreseeable disadvantages or risks in participating in this research. There is no compensation for taking part in this study. There are no individual benefits to taking part. By talking to us, you will contribute to creating more knowledge of how to design and implement feedback systems/dashboards in the best way so that it helps to improve neonatal hospital care in LMICs. This may help other implementers and elsewhere in the future, for example through learning from your experience and improving further implementations. We kindly ask you to participate in this study during a time outside your working shift(s).

7. What happens with the information provided and who will have access?

The information you provide during the study is the research data. Any research data from which you can be identified (name, current position/role, audio recording etc.) is known as personal data. Personal data will only be used by the research team to identify you for a follow up interview (as described above) and then be pseudonymised as appropriate. Pseudonymised data and other research data (including consent forms) will be stored for at least 3 years after publication or public release of the work of the research, while personal data such as for example your IP address will not be stored. We will take all reasonable measures to ensure that data remain confidential. Further information about your rights with respect to your personal data is available from https://compliance.admin.ox.ac.uk/individual-rights.

Research data will be transferred and stored securely on password protected computers, backed up on a server held by the University of Oxford in the UK. Identifiable data will be removed whenever possible and any data transfer will be done securely and with a similar level of data protection as required under the UK GDPR legislation. The principal investigator, supervisors and collaborators (named above) will have access to the research data. Responsible members of the University of Oxford may be given access to data for monitoring and/or audit of the research. The knowledge gained from this research will be shared in summary form, without revealing individuals' identities, with relevant stakeholders and the international research community through journal publications.

In future studies, we may use anonymised data collected during this study to support new research on feedback systems and dashboards by other researchers in other countries/settings. All personal information that could identify you will be removed or changed before information is shared with other researchers or results are made public.

8. Will the research be published?

The research will be written up as a student's thesis. On successful submission of the thesis, it will be deposited both in print and online in the University archives to facilitate its use in future research. If so, the thesis will be openly accessible. Furthermore, the research outputs may be published, e.g. as academic publications, websites. The University of Oxford is committed to the dissemination of its research for the benefit of society and the economy and, in support of this commitment, has established an online archive of research materials. This archive includes digital copies of student theses successfully submitted as part of a University of Oxford postgraduate degree programme. Holding the archive online gives easy access for researchers to the full text of freely available theses, thereby increasing the likely impact and use of that research.

9. Will I get reimbursed for participating?

You can ask for reimbursement of your data use (1GB or 250KES) if you fully complete the survey. However, reimbursement requires to share your phone number at the end of the survey. Your information will be kept confidential at any time.

10. Who has reviewed this study?

This project has been reviewed by, and received ethics clearance through, the University of Oxford. Ethical review of all procedures and data handling linked to the survey was conducted following the Oxford Tropical Research Ethics Committee (OxTREC) guidance (approval reference 532-21) and in line with the University of Oxford guidance on the requirement of data protection law.

11. Who do I contact if I have a concern about the study or wish to complain?

If you have a concern about any aspect of this study, please speak to the principal investigator or supervisor and we will do our best to answer your query. If you remain unhappy or wish to make a formal complaint, please contact the Research Ethics Committee at the University of Oxford who will seek to resolve the matter as soon as possible:

Principal investigator (student): Ms. Christiane Hagel, University of Oxford, Nuffield Department of Medicine, Centre for Tropical Medicine and Global Health, Peter Medawar Building for Pathogen Research, South Parks Road, Oxford, OX1 3SY, UK; Tel: +44 (0)7414 090265; Email: christiane.hagel@ndm.ox.ac.uk

Supervisor: Prof. Mike English, KEMRI Wellcome Trust Research Programme, P.O. Box 43640-00100, Nairobi; Tel: +254 (0)7226 28700; Email: MEnglish@kemri-wellcome.org

Oxford Tropical Research Ethics Committee (OxTREC), Research Services, University of Oxford, Wellington Square, Oxford, OX1 2JD; Tel: +44 (0)1865 282585 (Wednesday mornings, Thursdays and Fridays); Email: oxtrec@admin.ox.ac.uk

Appendix 5.7. Survey Analysis Plan

1. **Define the analysis plan**. What's the question? Choose the right method. Choose exposure and outcome, potential confounders (optional effect modifiers as well).
2. **Explore and understand the data**. What are the variables of interest?
   a. Continuous: Range, scale, distribution.
   b. Categorical: Number of levels, distribution, reference group definitions.
   c. Missing values: Assessed the extent, pattern, and how they were coded.
3. **Summarise the data**. Use step 2 to make sure 3 is succinct and informative.
   a. Descriptive statistics: Means, medians, standard deviation, proportions.
   b. Visual tools: Basic tabulations, histograms, box plots, bar charts.
   c. Table 1: Generate a summary of these summaries for primary exposure, outcome and possible confounders
4. **Data cleaning.** Ensure data quality, look for implausible / inconsistent values, outliers, errors, transformations.
   a. Then **RE-SUMMARISE** data after cleaning and assess changes.
   b. Document all steps to ensure transparency and reproducibility.
5. **Data reduction.** Choose variables for analysis.
   a. Use theory (incl. similar literature and validation studies) to guide selection.
   b. Evaluate the extent of missingness, spread, and informativeness of each variable. Consider dropping.
   c. Perform univariable analyses to explore associations with the exposure and outcome.
   d. Check for multicollinearity through correlation analysis. Check model fit later.
6. **Univariate analyses**. Describe each variable independently.
7. **Bivariate analyses.** Explore how two variables are associated with each other.
   a. t-test: Group comparisons of continuous outcomes.
   b. Chi-square test: Associations between categorical variables.
   c. Analysis of variance (ANOVA): Comparisons across more than two groups.
   d. Correlation: Linear relationships between continuous variables.
   e. Simple linear regression: One predictor, continuous outcome.
   f. Simple logistic regression: One predictor, binary outcome.
8. **Multivariable analyses** – are two or more variables associated with outcome?
   a. Multiple linear regression: For continuous outcomes.
   b. Multiple logistic regression: For binary outcomes.
   c. Cox regression: For time-to-event outcomes (if relevant).
   d. Generalised linear models: For count, ordinal, or other outcome types.

Appendix 5.8. Survey Questionnaire Structure

Survey pages	Theme	What was planned to get out of this survey	Key results
Page 1: Welcome			N=103 respondents
Page 2: Demo-graphics **Q 1 - 5**	**1**	Separate/stratify **people, target groups** by their demographics	Majority Age group 31 -40 (n=46 of 103, 44.7%) Gender female (n=69, 67%) Country Kenya (n=102, 99%) Position directly related to NN care: YES (m=69, 83.5%, NO n=5, 4.9% and PARTLY n=12, 11.7%)
Page 3-6: Context and tools in use **Q 6- 20**	**1**	▪ Separate/stratify people, target groups by their **experience/knowledge about using digital tools**, dashboards and CIN/NEST, and other tools in use ▪ Self-assessed feeling about **needs** related to tool in use	Main role: Clinical-hospital role (n=71, 68.9%) **Use of computer** for work in a) non-hospital context group: Every day100% **b) in clinical hospital role group: Every day around 60%** c) in non-clinical hospital role group: 100% **Use of smartphone** for work in a) non-hospital context group: Every day 90% **b) in clinical hospital role group: Every day 100%** c) in non-clinical hospital role group: 100% **Use of WhatsApp** for work in a) non-hospital context group: Every day 60% **b) in clinical hospital role group: Every day 85%** c) in non-clinical hospital role group: 20% Tools or approaches most used for feedback on performance are registries and summary forms, and HMIS (23.3% each), and staff meetings incl. visual presentations (22.3 %) which reflects on the Kenyan hospital context

			Usefulness rating: *to be adjusted to different groups by position (and maybe by age)* **Training needs rating**: respondent mainly responded that they need much more training related to how to use the tools, how to interpret data and how to use feedback to improve quality of care **Preference for other tools:** Digital platform that one can log on to, e.g. dashboard or website (n=56, 54.4%) **Preference of frequency** to receive feedback on Individual performance: Real-time 25%, Daily 15 %, Weekly 15%, Bi-weekly 4% Monthly 22%, Quarterly 14.7% Team's performance: weekly / monthly Hospital/organisation's performance: monthly/ quarterly
Page 7: Neonatal indicators **Q 21 - 24**	**2**	Assessment of whether **neonatal indicator needs and perceptions**	Almost all indicators that have been presented were rated between 6-7 (on a scale 1-7) as very important to receive feedback on, small number around 4% selected "irrelevant". Perception of main users of feedback on these indicators (multiple responses were possible): NBU staff 80%, Hospital Managers 63%, National/county Government 45%, HRIOs 39%, Researchers 34%
Page 8: CIN table **Q 25 -27**	**3**	**Overall satisfaction with CIN tool content:** Find out how different groups (CIN exposed, not exposed) perceive feedback report and frequency. This reflects more on the design and delivery of the tool	25: Table to identify if performance improved people seemed to have problems to interpret the table which is partly because of the question itself but also its seemed people did not know how to compare in general. 26.1- 26.5: Question if table is suitable in general, benchmarking, terminology, colour, figure legend majority rated mainly that it is very much suitable 27.1-27.2: Ease of interpretation and usefulness rating majority responded with very easy/useful despite having problems

Page 9 – 11: CIN graphs **Q 28 - 33**	**3**	**UX/ Usability of CIN visualisations** Specific to graph/form/delivery: ease of use, usefulness,	**28: Bar chart** n=23 of 103 correct answer and n=11 were close to correct answer as the bar chart is only a tiny bit bigger, so it is difficult to see there is a difference 29.1: Ease of interpretation rating: very easy with majority rating 6/7 29.2: Usefulness rating: very easy with majority rating 6/7 **30: Line chart** rating about ease of interpretation for different lines easy-very easy while line (around 30%) vs not easy at all around 7-10% 31: Rating entire visualisation a) easy to interpret information: easy-very easy around 23-29% vs not easy 7% b) usefulness of graph type: very useful40% vs. not useful 3% **32: Line chart with performance information lines/colours** Rating performance over time: 32a: correct improved 1-2: Answer correct for 1 13.6% and for 2 30.1%, “difficult to tell” 7%n and **35% were over interpreting performance improvement** 32b: correct improved a lot 3 and 63% correct, around 40% had difficulties 32c: correct declined a lot -2/-3: 42% correct, around 55% had difficulties 32d: In total correct probably small improvement 1 and 16.5% interpreted correctly 26% responded “difficult to tell”, around 55% had difficulties 33 Rating entire visualisation a) easy to interpret information: easy-very easy around 23-29% vs not easy 8.7% b) usefulness of graph type: very useful 48.5% vs. not useful 5%
Page 12: NEST **Q 34 - 37**	**4**	**UX/ Usability of NEST Dashboard** Top part of dashboard	34: 15%: correct response n=59 of 103 35: 172: correct response n=63 of 103 36 9%: correct n=80, n=6 said they don’t know what normal birth weight is 37: Rating a) Ease of interpretation: easy-very easy around 26-39% vs not easy 6.8% b) Usefulness: very useful 61.2% vs. not useful 1%

Page 13: NEST **Q 38 -44**	**4**	**UX/ Usability of NEST Dashboard** Middle part of dashboard	38: Prematurity and Infections: correct response for prematurity 80% and Infections 75.7% 39: None of the above: Correct 30% 40: Around 60: Correct 35.9% 41: Around 2: Correct 29.1% 42: Around 254: Correct n= 49/ 47.6% 43: Same q as 42, 254: n=28 44: Rating **a) Ease of interpretation 21% said not easy at all,** and only 15% said very easy, almost equal distribution between different response optione b) Usefulness: Very useful 41.7% vs not useful 6.8%
Page 14 NEST **Q 45 - 47**	**4**	**UX/ Usability of NEST Dashboard** Bottom part of dashboard	45: Catheters; correct response 68% 46: 48, correct 74.8% 47: Rating a) Ease of interpretation; Very easy 26-27 % vs not easy at all 10.7% b) Usefulness: Very useful 50.5% vs not useful 4.9.%
Page 15: CIN/NEST users **Q 48 - 50**	**1**	Stratify if respondent is used to work/ **exposed to CIN feedback report and/or NEST dashboard**	48: Used CIN before YES 78.8%, all knew what CIN is and find the report useful 97.6% Further written responses…. 149 Used NEST db before:YES 47.6% vs No 49.5%, don't know what it is 2.9 %, 87.7% found it very useful
Page 16: Confidence measure **Q 51**	**1**	**Confidence of respondents** in their role and related to quality of care improvement efforts	51 A to assess information B to use information C to have the skills D to be able to achieve improvements E to know what feedback on performance and quality of care means For all areas around 40-50% feel very confident

Chapter VI Qualitative Analysis Appendices

Appendix 6.1. Implementer Interview Tool

A. Objective of this interview tool: Understanding the perspective of implementers of feedback tools about their experiences when designing and implementing these tools, their usability and usefulness, and what challenges they experienced.

- Understanding the usability and usefulness of feedback tools from the perspective of implementer, their views on what works best and what the barriers and challenges were.

- The main focus is here on implementers to find out what their experiences are in terms of what guidelines they used to design the tools, how they selected KPIs and specific graph types, how they engaged with users and what their experiences were when designing and implementing these tools.

B. Introduction

- Hi, I am …
- **This study is about…**
    - …about hospital dashboards and feedback on performance tools. We hope to learn FROM YOU how you were involved in the design and implementation of the dashboard from your programme (CIN or NEST), your learnings from these processes and views on how it helped improving quality of hospital care for newborns.
    - …targeting implementers and (potential) users of these tools and so we hope to collect feedback from you and other people who were involved.

- **Can I crosscheck…**
    - …if you have read and understood everything from the consent form

- **Only if consent needs to be recorded** because signature could not be collected:
    - Can you please confirm that the study "The use of data to improve quality of neonatal hospital care in LMICs" was explained to you, that you have understood all that has been explained, and your questions were answered satisfactorily at this moment.
    - Can you please confirm that you understand that you can change your mind about your participation in this study at any stage and it will not affect you in any way to drop out.

- And do you now agree to take part in this research, that this online interview will be recorded and data be used as explained? And do you agree that we can make use of anonymous direct quotes from this interview?

- **Do you have any questions at this point?** *Expected questions could be:*
    - *Will you share the findings with me?*
      >>We will publish the results of this study in form of a PhD manuscript and research article. If you wish to receive a short summary, please feel free to reach out.
    - *How is this research linked to other KEMRI studies I've seen around here?*
      >>This research is part of and linked to CIN and NEST360
    - *Will what you've learned from these interviews lead to change?*
      >>We hope the results of these interviews will help us to improve the usability and usefulness of feedback reports and therefore help to improve quality of care for newborns. So yes, we hope this will lead to a positive change.

Notes on what observations to make during an interview:

- Take notes of the implementers' language and confidence when referring to feedback tools, their content and how they refer to users' interactions with it.

Appendix 6.2. User Interview Tool

<u>A. Objective of this tool:</u> Understanding the usability and usefulness of feedback tools for users and their needs and preferences, how these might be met, barriers and contextual factors.

- Understanding the usability and usefulness of feedback tools for users and their needs and preferences, how these might be met, barriers and contextual factors.

- The main focus is here on (potential) end-users to find out what their information needs are, both in terms of which content/KPIs they need to see and what kind of design/graph type they would prefer so that it is easier to understand, why/which different elements from dashboards/feedback report work better than others.

B. Introduction

- Hi, I am …
- **This study is about…**
  - …about hospital dashboards and feedback on performance tools. We hope to learn FROM YOU how we can design these tools better so that they meet your needs and so they help improving quality of hospital care for newborns.
  - …in collaboration with the Clinical Information Network (CIN) and NEST360
  - …targeting (potential) users of these tools and so we hope to collect feedback from paediatricians, nurses and HRIOs/data clerks from different hospitals. We identified you as one of the (potential) users of these tools and so we hope to learn from your experience.

- **Can I crosscheck…**
  - …if you have read and understood everything from the consent form
  - …if you filled out the online survey on dashboards during one of the CIN meetings on 5/6 May (Nairobi Royal Tulip) or 9 of May (CIN pre-KPA meeting in Mombasa)?

- **Do you have any questions at this point?** *Expected questions could be:*
  - *Will you share the findings with me?*
    >>We will publish the results of this study in form of a research article latest around the beginning of next year. If you wish to receive a short summary please feel free to reach out.
  - *How is this research linked to other KEMRI studies I've seen around here?*
    >>This research is part of and linked to CIN and NEST360
  - *Will what you've learned from these interviews lead to change?*

>>We hope the results of these interviews will help us to improve the usability and usefulness of feedback reports and therefore help to improve quality of care for newborns. So yes, we hope this will lead to a positive change.

Notes on what observations to make during an interview:

- Take notes of the user's language and confidence when referring to feedback tools, their content and visualisations.
- Observe what components/ visualisations of tools the user highlights.

- **Only if consent needs to be recorded** because signature could not be collected:
    - Can you please confirm that the study "The use of data to improve quality of neonatal hospital care in LMICs" was explained to you, that you have understood all that has been explained, and your questions were answered satisfactorily at this moment.
    - Can you please confirm that you understand that you can change your mind about your participation in this study at any stage and it will not affect you in any way to drop out.
    - And do you now agree to take part in this research, that this online interview will be recorded and data be used as explained? And do you agree that we can make use of anonymous direct quotes from this interview?

Appendix 6.3. Implementer Interview Participants

Interview #	Gender	Position	CIN / NEST implementer	Country focus	Other notes
1	F	Clinician, senior researcher	NEST	Kenya, Malawi, Tanzania and Nigeria	
2	F	Study coordinator	CIN & NEST	Kenya	
3	M	Data manager and health information researcher	NEST	Kenya, Malawi, Tanzania and Nigeria	
4	F	Study coordinator, researcher	NEST	Kenya, Malawi, Tanzania and Nigeria	
5	F	Study coordinator, researcher	NEST	Kenya, Malawi, Tanzania and Nigeria	
6	M	Study coordinator	CIN & NEST	Kenya	
7	F	Study coordinator	CIN & NEST	Kenya	
8	M	Clinician, senior researcher	CIN	Kenya	
9	M	Data manager and health information researcher	CIN	Kenya	
10	M	Data manager and health information researcher	CIN	Kenya	
11	M	Country representatives/ people on the ground	NEST	Malawi	Involved in implementation on the ground
12	F	Country representatives/ people on the ground	NEST	Tanzania	Involved in implementation on the ground
13	M	Clinician, senior researcher	CIN	Kenya	CIN tool user as clinician

14	M	Data manager and health information researcher	CIN	Kenya	
15	F	Study coordinator, researcher	CIN	Kenya	CIN tool user as clinician
16	M	Data manager and health information researcher	CIN	Kenya	
17	M	Data manager and health information researcher	CIN	Kenya	
18	M	Study coordinator	CIN & NEST	Kenya	
19	F	Study coordinator, researcher	CIN	Kenya	
20	M	Data manager and health information researcher	NEST	Kenya, Malawi, Tanzania and Nigeria	
21	M	Country representatives/ people on the ground	NEST	Nigeria	Involved in implementation on the ground
22	F	Study coordinator, researcher	NEST	Kenya, Malawi, Tanzania and Nigeria	

Appendix 6.4. User Interview Participants

Interview #	Gender	Position	Exposed to CIN and/or NEST	Participation in survey
1	M	Nurse	CIN & NEST	-
2	M	Nurse	CIN & NEST	-
3	M	HRIO	CIN	+
4	F	Data clerk	CIN	+
5	F	Nurse	CIN & NEST	-
6	F	HRIO	CIN & NEST	-
7	F	Data clerk	CIN	-
8	F	Nurse	CIN & NEST	+
9	F	HRIO	CIN	-
10	F	HRIO	CIN	+
11	F	Nurse NBU	CIN	-
12	F	Nurse NBU	CIN	-
13	F	Paediatrician	CIN	+
14	F	Paediatrician	CIN & NEST	+
15	M	Nurse NBU	CIN & NEST	+
16	M	HRIO	CIN & NEST	Unclear
17	F	Paediatrician	CIN	-
18	F	Paediatrician	CIN & NEST	+
19	M	HRIO	CIN & NEST	+
20	F	Nurse NBU	CIN	+

*All nurses were wither Deputy nurses in charge or Chief Nursing Officer,
**HRIOs are official hospital employees but data clerks are linked to the project team as team members but working in the hospitals

Appendix 6.5. Coding Frameworks

6.5.1. Coding Framework Implementer Interviews

Implementer Themes	**Sub-themes**	***Inductive coding***
1. Designing content and form	Design steps/process	
	Best practices	*Informed by guidelines*
		Informed by individual knowledge/expertise
	Data	*Data (sources, collection analysis)*
		Data quality (missing, standardised, timeliness etc.)
	KPI priorities	
	Visualisation	*Graph types*
		Colour coding
	Benchmarking	
	Other key considerations	*NEST specific*
		CIN specific
2. Implementing the tool & processes	Delivery	*Information delivery with tool*
		User access to the tool
	Evaluating design	*Iterative/ dynamic/ continuous*
	User engagement	*Co-design*
		User feedback
		Other engagement

	User challenges	*Workload*
	Usability	*Ease of use, Usefulness*
	User reaction	*Acceptance*
		Accountability
		Motivation
		Trust
	Difference backend frontend	
	Other key considerations	*Challenges*
		Best practices
		What worked well
3. People being involved	Interviewee details	*Role*
	Other tool users	*Clinicians*
		Hospital management/leaders
	Other contributors	*Data collectors*
		QI team
		Facilitators
	Researchers	
	MoH	
	Other key considerations	
4. Engaging with the wider system	Integration to existing Health Information System or workflow	
	Other forms or tools	
	Complexity of the system	
	Other key considerations	
5. Making an impact	Purpose of the tool	*Making sense of data*

		Use of data, information to action, change
		Action plan, actionability
		Decision-making
		Identify gaps or challenges
		Behaviour change
		Performance, quality improvement
		Empower people/ facilities
		Teamwork
	Future potential	*Alert-system*
		Scale of use
		Sustainability
		Real-time
6. What is a dashboard?		

6.5.2. Coding Framework User Interviews

User Themes	Sub Codes	Examples of information collected
1. Design of content and form	General	What users look at first, KPIs of interest, Benchmarking, Preference of design, Usefulness, Ease of Use, Engagement in the design or implementation of the tool
	CIN	
	NEST	
2. Implementation,	General	

	user needs and preferences		Time available/restrictions to use the tool, Preference frequency of delivery, Preference form of delivery, Training needs and other needs
		CIN	
		NEST	
3.	**People/ Users being involved**	General	Interviewee details
		CIN	(Frequency of) interaction with tool, Main/ secondary user, other people being involved e.g. midwife,
		NEST	
4.	**The wider system**	General	Challenges within system that affect purpose/goals/tool, other tools or forms of feedback/ non-feedback, Data sources, collection, quality, TRUST
		CIN	
		NEST	
5.	**Achievement of purpose and goals with dashboard**	General	Acceptance/ Accountability/ Owning tool or data, assistance in achieving goal, expectations of achievements in role, motivation, knowledge, using information to improve data quality/ performance/ quality of care, relationships/ interactions with others (QI, colleagues, gov), other takeaways from the tool
		CIN	
		NEST	
6.	**What is a dashboard?**		

Appendix 6.6. Critical Appraisal Skills Programme (CASP) Checklist

CASP Checklist Question [138]	**Answer** Yes/No/Can't tell	**Notes**
Section A: Are the results valid?		
1. Was there a clear statement of the aims of the research?	Yes	The aim of this research was clearly defined to understand the impact of dashboards for ensuring data is used for quality of care improvements. The importance of this study was established through an extensive review of audit & feedback and design, a scoping review and local stakeholder consultations.
2. Is a qualitative methodology appropriate?	Yes	The qualitative component of the mixed-method study was chosen to deeply explore the subjective experiences of dashboard users and implementers to add these narratives, which quantitative methods could not adequately capture. As such adding a qualitative component was appropriate to answer the research questions.
3. Was the research design appropriate to address the aims of the research?	Yes	The research design included semi-structured interviews, which were justified as they allowed for in-depth discussions and insights that align with the study's aims.
4. Was the recruitment strategy appropriate to the aims of the research?	Yes	Participants were selected based on their direct or potential experience with using the CIN and/or NEST dashboard. This purposive sampling ensured that the most relevant insights were gathered.
5. Was the data collected in a way that addressed the research issue?	Yes	Data was collected through semi-structured interviews conducted online (implementers) during the COVID-19 pandemic, and in person (users) in a convenient setting and convenient time for participants. Detailed notes and recordings were taken to ensure comprehensive data collection.

6. Has the relationship between researcher and participants been adequately considered?	Yes	Reflexivity was practiced by taking notes and discussing early findings with the supervisory team and research assistant to reflect on any potential biases and influences. Transparency in my role and the research process was ensured throughout.
Section B: What are the results?		
7. Have ethical issues been taken into consideration?	Yes	Ethical approval was obtained from the relevant ethics committee at the University of Oxford, see OxTREC proposal in Appendix 3.3. Participants were fully informed about the study, consent was obtained, and confidentiality was maintained.
8. Was the data analysis sufficiently rigorous?	Yes	A detailed and systematic analysis process was followed, including coding, thematic analysis, and supervisor debriefing to ensure reliability and validity of the findings.
9. Is there a clear statement of findings?	Yes	Findings were explicitly stated, with supporting quotes from participants. A balanced discussion of the evidence for and against my arguments is provided, ensuring a credible presentation of results.
Section C: Will the results help locally?		
10. How valuable is the research?	Yes, it is valuable.	The research contributes valuable insights into the design and implementation of dashboards to translate data into useful information that aims to improve hospital performance and as such quality of neonatal care. The study identified areas for improvement and providing recommendations for policy and practice. The findings can inform similar initiatives in other contexts.

Appendix 6.7. Consolidated Criteria for Reporting Qualitative Research (COREQ) Checklist

Adapted Checklist from Tong, Sainsbury, & Craig, 2007, [137]

Item No.	Topic	Guide Questions/Description	Description or Link
Domain 1: Research team and reflexivity			
Personal Characteristics			
1	Interviewer/facilitator	Which author(s) conducted the interview or focus group?	CH & JJ, Section
2	Credentials	What were the researcher's credentials? (e.g. PhD, MD)	Master's degree
3	Occupation	What was their occupation at the time of the study?	Doctoral Researcher & Research Assistant
4	Gender	Was the researcher male or female?	Female
5	Experience and training	What experience or training did the researcher have?	Section 3.5.
Relationship with participants			
6	Relationship established	Was a relationship established prior to study commencement?	Section 3.4.
7	Participant knowledge of the interviewer	What did the participants know about the researcher? (e.g. personal goals, reasons for doing the research)	Section 3.4. and Appendix 5.6.
8	Interviewer characteristics	What characteristics were reported about the interviewer/facilitator? (e.g. bias, assumptions, reasons and interests in the research topic)	Section 3.5.
Domain 2: Study design			
Theoretical framework			

9	Methodological orientation and Theory	What methodological orientation was stated to underpin the study? (e.g. grounded theory, discourse analysis, phenomenology)	Section 3.2.
Participant selection			
10	Sampling	How were participants selected? (e.g. purposive, convenience, snowball)	Section 6.2.
11	Method of approach	How were participants approached? (e.g. face to face, telephone, mail, email)	Section 6.2.
12	Sample size	How many participants were in the study?	42, see Section 6.2.
13	Non-participation	How many people refused to participate or dropped out? Reasons?	None
Setting			
14	Setting of data collection	Where was the data collected? (e.g. home, clinic, workplace)	Section 6.2.
15	Presence of non-participants	Was anyone else present besides the participants and researchers?	Only CH or JJ
16	Description of sample	What are the important characteristics of the sample? (e.g. demographic data, date)	Appendix 6.3. Appendix 6.4. Section 6.2.
Data collection			
17	Interview guide	Were questions, prompts, guides provided by the authors? Was it pilot tested?	Yes, Section 6.2., Appendix 6.1. and Appendix 6.2.
18	Repeat interviews	Were repeat interviews carried out? If yes, how many?	No
19	Audio/visual recording	Did the research use audio or visual recording to collect the data?	In-person interviews audio, virtual with visual recording
20	Field notes	Were field notes made during and/or after the interview or focus group?	Yes, during and after
21	Duration	What was the duration of the interviews or focus group?	About 1h

22	Data saturation	Was data saturation discussed?	Yes, CH, JJ, SM and ME
23	Transcripts returned	Were transcripts returned to participants for comment and/or correction?	No
Domain 3: Analysis and findings			
Data analysis			
24	Number of data coders	How many data coders coded the data?	One (CH)
25	Description of the coding tree	Did authors provide a description of the coding tree?	Appendix 6.5.
26	Derivation of themes	Were themes identified in advance or derived from the data?	Inductive and deductive, Section 6.2.
27	Software	What software, if applicable, was used to manage the data?	MAXQDA
28	Participant checking	Did participants provide feedback on the findings?	Conferences, Appendix 7.2.
Reporting			
29	Quotations presented	Were participant quotations presented to illustrate the themes/findings? Was each quotation identified? (e.g. participant number)	Section 6.3. Appendix 6.8. Appendix 6.9. Appendix 6.10.
30	Data and findings consistent	Was there consistency between the data presented and the findings?	Yes, linked to people, tool a systems see Table 14, Table 15 and Table 17
31	Clarity of major themes	Were major themes clearly presented in the findings?	Yes, seen in Table 15 and Table 17
32	Clarity of minor themes	Is there a description of diverse cases or discussion of minor themes?	Yes, seen in Table 15 and Table 17, and Appendix 6.5.

Appendix 6.8. Quotes: People Involved with Dashboards, Section 3.1.

Quote #	Theme	Respondent	Quote
Implementer perspectives of roles as part of being involved with the CIN and NEST dashboards			
1	Data management	CIN Implementer Interview 14 Line 5	*"So, my role in CIN, I am the senior data manager. I lead the data management team. It's a relatively big team and that team is composed of data managers and assistant research officers and some statisticians. My role really is to make sure that the data - I mean, I put in place mechanisms on the systems which now see the data collection and transmission of that data from the hospital side to our central server. That, really setting up of that architecture from creating the REDCap forms where that data can be entered, making sure that the data clerks are being trained properly so that then they can do that extraction of data from medical files into our systems. That seamless transfer of data. Then once the data has been transferred to a data clerk machine then again, creating a synchronisation module which will make sure that that data is safely again transmitted from that hospital to our end. Making sure that end-to-end, there's encryption and that in between here, there is no third-party intrusion."*
2	Data quality and interactions with users	CIN Implementer Interview 15, line 5-7	*"I was the study coordinator for the Clinical Information Network and I basically would ensure quality of data and the interaction between the data team, the clinical team and feedback to the hospital team…. This was through physical meetings, quite a number of physical meetings that we used to have. We had WhatsApp groups. We will get ideas from the clinical team on what they are interested in knowing and*

			then the same will be discussed but with the hospital team. The same will be discussed with the PIs in the clinical team in the office."
3	Data quality assurance	CIN Implementer Interview 13 Line 5	*"I co-lead the management of the CIN. My background role is to manage the team that does the data quality assurance, and to review the data. The data quality on a weekly basis, to make sure that this data is of good quality. Then when the reports are due, we have a look at the reports to make sure that there are no errors in the reports and that sort of thing before sending out to the reports, so in a nutshell, that's my role."*
4	Tool development	CIN Implementer Interview 17 Line 7	*"I am actively engaged in the development of creating tools, which basically I use to do the reporting, and the generation of reports, which I usually share with the hospitals, and the ministries, and other stakeholders. I also take part in handling requests, data requests, which involve the use of CIN* data. Yes, that's much of what I do."*
5	Tool development	CIN Implementer Interview 10 Line 6	*"I'm involved in both dimensions of collecting the data in terms of looking and supervising the people who collect the data, because that's the beginning point. Then when the data is with us and we have checked on the quality, I'm involved in the design of reports. We have reports that have been there even before I came to KEMRI* and we have the reports that we have designed when I've been here, and I've been involved in from the word go. My role has been to receive opinions or suggestions of what should be included from them, and script that and implement it, give it to them, take them through, help them understand the logics involved. Once that has been confirmed and agreed upon, then we decide on the frequency of sharing these reports."*
6	Data quality improvement	NEST Implementer	*"As part of NEST I look at the more health systems side of things, and so it was pretty much to try, try to get other pieces of the data coming in and other aspects that need to be tracked, just to see how best to*

		Interview 6 Line 20	*structure that into a dashboard. Then having said that, <u>this was part of a huge team that put together a dashboard</u>. I think a lot of my inputs are around the HR, quality of care because that's where a lot of my past work has focused on, and inputs around how best to present some of that data, obviously with inputs from others. <u>I think my role going forward is trying to see how the dashboard is being used as part of data feedback, pretty much linked to quality improvement</u> that are happening in hospitals and tracking how those are… So how dashboards have been, first been fed back to hospitals, and how those clinical teams within hospitals are using that data to influence change and track that over time."*
7	Data collectors	NEST Implementer Interview 20 Line 41 - 49	*"I think there has been concerns around, I think actually there's been an underestimated patient [unclear word 0:23:39.8] of the level of workload involved in the data collection to generate the dashboard. It's not really output, but I do think it's important from the experiences that I've had working with all the data managers in the country, <u>I do think it's important that each facility has its own data collector.</u> At least one, if not two." …* *"For instance, in Malawi, we have <u>data collectors that are sort of government.</u> In Nigeria, we have data collectors that are employed by [unclear 0:26:51.8]. I would say that depends on who's collecting, probably yes, I would imagine. In Malawi, for instance, it would be interesting to see the differences between Malawi and maybe Nigeria, for instance, I would say, because the <u>data collectors are employed by different people.</u> I would be guessing but I would say there's probably a difference."*
8	KPI selection	NEST Implementer Interview 12 Line 7	*"Yes, my contribution was especially on the indicators, like which <u>indicators should appear in the dashboard</u>. When we piloted it, we see like, you know - we had several comments from the facilities, like we need maybe more disaggregation, like [for outborns, inborns 0:03:14.7], showing these indicators. Why don't you put these in that, but the active part of it, I'm not - but at least for the indicators in the analysis, yes, that one, I participated very well."*

User perspectives about other important actors for using data to improve quality of care			
9	Champions	Paediatrician, User Interview 13 Line 18	*"So, then we would have a meeting usually we have scheduled mortality meetings, mortality or morbidity. And we would go through the highlights of the report rather than the entire report. So, what I do is I receive the report, a few others in my team also receive it, the nursing officer in charge or the CIN champions, and we synthesize because it's usually not possible to go through the report step by step with the whole team. And once we synthesize, we look at the key highlights and then discuss those with the team. This discussion comes into two forms. One, we have formal meetings like in the mortality meetings have mentioned our CMES, in case there are issues that are crosscutting. But more importantly, there are things that I personally use in the day-to-day delivery of service."*
10	Benchmarking & Collaborations	Paediatrician, User Interview 18, Line 25 & Line 42	*"I would want to know the reasons for their successes and really looking at their good side like what is it that you guys are doing really great and so I would contact my paediatrician, my fellow paediatrician like "we are seeing that you guys are doing really well, what is it? How did you get to that point?" And then we can communicate and say my administration has done this for me, we've gotten equipment maybe, we've gotten more staffing ratios, or improved staffing ratios or improved maybe continuous medical education, we are doing it every month, or every week, we are doing audits and all those things. Or we have partners who are helping us with funding and all those and so those are important things that I feel can be of assistance." […] "When you compare with other hospitals there is a bit of pressure that comes with it, like administratively obviously. People in the offices are like what are you guys doing wrong? Why are we at the bottom and what is happening? But it would be a good thing because if you were trying to be heard before, now you have the data, you have the evidence to say look we are doing badly because we don't have staff, can you do something about this otherwise I will always be at the bottom. So, as much as it will take a hit on our ego, but it will assist in giving people evidence that they require to make action,*

			because administrative work, you need evidence, you need proof and you don't just come and tell me we need staff, I will tell you the reasons why."
11	Government as data user	Paediatrician, User Interview 19 Line 6	*"You see the way our government works is that they make decisions based on their interests but not based on the data. You see you can issue data that this is what we have but when they make a decision it's totally different from what we are doing depending on the interests of the managers. But I would wish that we make decisions based on what we are collecting like this data we are collecting here. Of which at times it is used, at times not used."*
Implementer perceptions on the primary and secondary users of dashboards			
12	The range of users and target groups	CIN Implementer Interview 16 Line 14	*"Main users include the <u>government, stakeholders, the hospital management and the people who are medically working with the patients, that is doctors, clinicians and nurses.</u> So, the same also is applied to the wider team which we work with across the department and the programme."*
13	Clinical frontline workers as primary users	CIN Implementer, Interview 8 Line 15	*"We like to think that the <u>hospitals staff are the main users.</u> We consider that the <u>paediatrician and some of the senior nurses</u> have a sort of pivotal role in trying to shape the practices of these quite junior clinical staff. So, most of the people that see the babies or see the children are <u>medical interns.</u> They're just fresh graduates. In theory they should be under the supervision of these more senior people, so we direct most of our feedback to these more senior people in the hope that they will then coach and supervise the junior people to do things better."*
14	Staff turnover as the reason	CIN Implementer	*"And there's two reasons for that. One is because that is part to really in one sense we can sort part of their role, part of their responsibility as the senior people in on these clinical units. The other thing I guess*

	for targeting specific primary user groups	Interview 8 Line 15	*it's much easier to feedback to a couple of key individuals than it is to try and keep feedback to everybody. And particularly because in the wards in which we focus the junior people who see the patients, completely change every three months. If we were doing feedback to the individuals who saw the patients, we'd have to identify who that was every three months, find a way of feeding back to them. So, the more senior people are much more, they stay in post much longer, so it's a combination of being practical and because they should have that responsibility.*
15	Main user & secondary user	CIN Implementer Interview 13 Line 18	*"Okay, so the main user, the intention of the main user being the hospital team rather than the individual. An understanding of the team, at that point, is the clinical team, so that includes the nursing officers, and also the doctors who are the primary care givers at the bedside, but also understanding that there are other caregivers, which include nutritionists and that kind of thing so anyone within the team, would usually be involved in the caring for the unit, that would be the primary user. [..] At the hospital level, a secondary user would now be the hospital administration when the team leaders from the units, that includes the nursing officers and the paediatricians when they are attending hospital meetings, where decisions are being made about resource allocation and people are bargaining for resources, then they use that information to back up their requests for their needs in their unit. That does allow the hospital administration to make rational decisions based on, that is demonstrated in the data."*
16	Other users expected to benefit from data	CIN Implementer Interview 8 Line 15	*"And then there is a whole set of users who use these data for aggregate reporting and publications. There's a whole set of users who benefit from this for their research careers. And ideally there's a set of stakeholders who would benefit, who should benefit, from these data from across multiple hospitals to know what's going on in hospitals and that they're the people at the more Ministry of Health."*

17	NEST confirming staff shortage	NEST Interview 11 Line 57	*"We expected that we would be meeting all the hospital management team during the QI visit to discuss the issues on the dashboard. When we say the hospital management teams, we are talking of the hospital director, the hospital matron, the hospital district nursing officer, the hospital medical officer, the pharmacy manager. All those people, we expected them to be there and discuss the issues on the dashboard, but when we started conducted the visits on the ground, we noted that those people are always busy."*
18	NEST user groups	NEST Interview 6 Line 22-26	*"From the dashboard perspective, I would say there are maybe four main groups. One is obviously what we call the implementation team, which is what you would also call the researchers, so the implementation team mainly being people within NEST who are looking at the data, because they need to track the data anyway as part of the intervention, tracking and process evaluation." […] "the second main group are users at hospitals, so at a hospital there is obviously the data or the health information team at the hospital (a), because a lot of data is their responsibility. They're the ones who are generating this." … I think the other team is the clinical team (b), because there's quite an amount of data that is linked to quality of care and how that is being tracked and seen." … "I think the other group at the hospital is the more managerial administrative team (c)" […] "Then, I think the last one at the hospital are the technical biomedical teams (d), who are responsible for looking after the medical devices."*
19	Parents as potential users	*NEST Implementer Interview 5 Line 17*	*Moderator: "What about patients or parents? Could that be a user group at any point in the future?" Respondent 5: "Interesting question. Yeah, I guess for the NEST dashboard I don't really see that being terribly useful because it's more about the neonatal unit overall. Yeah, it's more sort of like a population level of the neonatal unit, not really for an individual parent or a patient."*

Appendix 6.9. Quotes: Dashboard Design and Implementation, Section 3.2.

#	Theme	Respondent	Quote
Implementer Perspectives: Varied Responses about Dashboard Co-design with Users			
20	"It was not co-designed"	NEST Interview 12 Line 27	*"No, they didn't have any chance."*
21		Interview 4 Line 48	*"I don't think it was, no. I think it was made by people who made dashboards in the past. I think there was also a big-time crunch, when it was first made - just what I've heard; I wasn't a part of it - that I think happened really fast. They were like, 'We need this now!' No, I don't think there was a co-design process."*
22	"It was co-designed"	NEST Interview 20 Line 16	*"I would say it was co-designed. I think the real co-design came before I started actually. I would think that would be the most, I would suggest that, I wasn't there but I understand a lot of the co-design was done with PIs etc., on the ground before that, that's my assumption. Actually, from there forward, I think it has been a constant reiterative process between talking to people and nurses, doctors, and biomed technicians withinside NEST and outside, to try and bounce ideas off and there has been this constant process of updating and improving. I also, just to reflect on that, another real reflection on the design would be that, for people, like C* for instance, I know that when she's gone to facilities to implement, there's been a real-life feedback process within that. I think that's an area that we really now, reflecting on it more, I would say an area where we really do get a lot of subsequent points to change things. I think a lot of feedback is coming through that route as well. Probably more so than others actually, now I think about."*

23		NEST Interview 5 Line 22	*"Yeah, definitely co-designed I think from that initial meeting up in Nairobi. But before I started, there were maybe like 30 or 40 key stakeholders from NEST and other partners that were there. And they proposed initial items, but I think even all the sort of changes that have been made to the dashboard have been very collaborative and co-designed across many different teams."*
24	Co-design challenge	NEST Interview 1 Line 30	*I think everybody says that everybody says co-design. I think the problem is who are you co-designing with? So the reality is the middle part of the dashboard that people are most passionate about it is the clinicians, so someone has someone who was a clinician (?). Still, I'm a clinician at heart. You know, I see their passion, but it's impossible because what happens is they want 100 different clinical things that actually aren't measurable. So they want all. At the denominators to be very clinical case definitions that have been very poorly done and inconsistent, especially in registers and notes. And so you need to be very aware with. So in in NEST and co-design we have four countries and four very different constituencies. You know what's the biomed want there? They want their device. What does the clinician want? They want a complicated, you know clinical definition with lots of nitty gritty "If this, then that" sort of thing whereas data person want something that's clear and can be measured. And what does the quality improvement person want? A different person want a different thing every time. So that's actually not helpful. So co- design is important, but it's one of those things that people just say. But how do you do co-design across four countries with four different groups? So what you need to do is you need to set some parameters that everybody agrees to. So the way with the clinicians to do that was to say we've only got one page. You can in this page, in this part you can have four pathways. Therefore you can't have all of them now. Which ones do you think people can change the most? So for example, management of HIV is really important, but what big things can the clinician do? a lot less than the obstetrician I'm afraid, it should be much more about prevention. So in that case you know which ones here… so congenital abnormalities really important, but at the moment if we haven't got a big stuff on pediatric surgery. There's less we can do here, so you know, I, I think it's easy to say co-design. What you have to do is to recognize that in co-design, you have four groups and then you also have to recognize power imbalances between them. So the clinicians, especially the doctors and the*

			senior ones have much more power. The people, interestingly, who had the least power, were the data people. And other people have to run the dashboard. And then if the data people don't understand the clinical definitions, that's really a problem. So then you need people who understand both clinical and data. So I think you know what I found was that if you didn't have people who, in the co-design process, who were able to help one person to see another person inside, that the co-design just doesn't work and you need to be very careful with this because you have to, you know at the end of the day the data people have to be able to manage the data and they have the least power. So if the co-design is clinicians only then the whole thing will fold, which is actually exactly what's happened with quite a lot of clinical dashboards."
25	"Piggy-backing" on CIN	NEST Interview 22 Line 105	*"Most of the information on the dashboard comes from the routine data systems, so the registers and neonatal forms. We do have those health systems inputs that come from the previous QI visits. The QI visits are not about data collection but some data is collected. For example, the functionality of the devices on the ward that day, or the number of staff present, or the number of babies to calculate bed occupancy rate; that kind of thing but, overall, minimal. Now, in Malawi - so, different countries are also in different stages of this. In Kenya, this all directly links with the CIN, piggy-backing on that data system. The last thing we want to do is have another parallel system." –*
Implementer perspectives: Challenges with Dashboard Development			
26	Data quality	CIN Interview 9 Line 26	*"First of all is <u>to ensure that the data is high quality, that we have collected all the data so that we have not underreported because the first thing that clinicians, the hospital team will look at is if the numbers</u> tally with what they… Before the reports are generated, we first have to ensure that all files have been captured at the end of every month. Then we can confirm that the reports are now ready to be save...*

			I also get any feedback. By feedback I mean anything that could be of concern from the people giving the feedback so that we can know how to improve the reports in the next feedback."
27	Visual presentation	CIN Interview 10 Line 6	*"Then when the data is with us and we have checked on the quality, I'm involved in the design of reports. We have reports that have been there even before I came to KEMRI* and we have the reports that we have designed when I've been here, and I've been involved in from the word go. My role has been to receive opinions or suggestions of what should be included from them, and actually script that and implement it, give it to them, take them through, help them understand the logics involved. Once that has been confirmed and agreed upon, then we decide on the frequency of sharing these reports."*
28	Decisions on how to visualise	CIN Interview 13 Line 28	*"I think, this is purely opinion, but I think that largely reflected the background of the CIN* teams, so lots of those decisions were made similar to how you'd make a decision when you're preparing a manuscript for publication. What should I have, a figure for what should I have, so I think a lot of that was informed from a research point of view of how best to present different sorts of data?"*
29	Integration	NEST Interview 1 Line 40	*"…the problem is they just want to go faster and faster. Like X [hospital contact person] they are like you know, the maternal guys have got a dashboard that we going to link up, you know you guys need to give it to everybody now electronically….So, I mean there's definitely demand and interest. I think that particularly when people start to have it electronically and look at it and use it regularly, I think that'll be the real power of it. … I think the first time there was a workshop in Kenya…I think the first time people were a bit confused in Kenya whether it was kind of a project. That's another thing about who's doing it. I don't understand who you're fighting with. You know it was. It's a sad show at the moment there. But the last round where it was really the government leading it and people getting excited that it was just completely different. They went to a facility and the facility would be totally excited about it. So, I think the politics in the team kind of took away from what the facility wants. Which is a sad indictment on us. We should be serving facilities, not fighting with each other."*

30	Integration to DHIS2	CIN Interview 13 Line 101	*"I think we need to get the principals right from the paper-based, and if we find a way of making that integration happen right now, so you can see how they complete the cycle of the whole day, from the point of patient evaluation to the movement of that information into the DHIS2. Then when the health records, the EHR start coming in, then people will already have an understanding of what the system requires to do. The software is the only platform that implements a process that is really clear. It is important because even for sustainability and also in the sense that since CIN* cannot be in every single hospital in the country, there are bits like, for instance, if we have the registers as a basic tool in each facility, then how does that information, that basic information from the registers, for instance, get into it into DHIS2. I don't know if that's the right time for it, a modular form of integration, but the best case scenario, all these sets of information can make it into DHIS2. This is how it would work, but as a minimum, I'm thinking of the registers as a starting point, which is something which can happen in all new-born units, even those ones which are not necessarily in CIN."*
31	Formal training	CIN/NEST Interview 7 Line 115	*"We didn't do any formal training to our teams on the dashboard, so as we were introducing Quality Improvement, we came with the dashboard and then I will tell them, okay, here is the dashboard and this is what it means, so when you get [signal breaks up] this is what it means. This colour means this, and this shows X, Y, Z, so I wish, I think we would have done it better so that maybe people understand what the dashboard is all about. Then maybe train the main focal persons or leaders maybe of these tracks so that maybe, so that they can tell the other people in their team so that everybody is in the know."*
Implementer Perspectives: Potential Unintended Consequences of Concern			

32	Demotivation	CIN Interview 19 Line 63	*"The mortality rate was just constant….It made them feel helpless and discouraged, so they're like, 'There's nothing we can do to improve our mortality rates because, I mean, that's just… I mean, what else can we do? We're a public hospital. We don't have an NICU."*
33	4-layer traffic lights systems	CIN Interview 8 Line 21	*"For the design it wasn't a particularly in-depth approach to looking at what was good practice. You know we have adopted largely a traffic light system, and which seems to be very commonly used, and therefore we I guess we felt it was intuitive. We slightly adapted it so most traffic light systems have three-layer levels, and we introduced a fourth, so that we could have a little bit more of a… Well, what we hoped is something that was <u>a little bit more discretion</u> in the middle zone. We have essentially a category of very poor, a fair which is just above very poor, good and then very good. So the aim is to try and use those like all traffic light system is. The aim is to try and that people will be aiming to try and increase their performance against those measures. And so it would be trying to achieve very good, so that's one aspect and therefore we use the standard colour coding for those kind of things. And the other thing that we had learned from earlier feedback literature that we had read was about the issue of benchmarking or comparison."*
34	Risk of media outlet & politics	CIN Interview 19 Line 99 ff.	*"I think one of the things maybe is if this information is interpreted by the wrong person, especially things like mortality rate, it can be used to castigate the newborn unit team. Yes, so if there's politics involved in the hospital, it can get in the wrong hands, go to the county and the end result is that this hospital is not doing anything. They're killing the babies. We've experienced that. Just politics, but not necessarily because it's true, but just because the report was interpreted in a political way to gain political mileage, again because there's a lot of politics in the counties. … Also, when a media outlet wants to report something about maybe our newborn unit, that's the details we provide them with, so someone can… If we don't know what our mortality rate is, that's the data we'll provide them with, so now it depends on how they spin the story and which politician reads it and decides to use it against us."*

35	Increased data collection workload	NEST Interview 20 Line 44	*I think there has been concerns around, I think actually there's been an underestimated patient [unclear word 0:23:39.8] of the level of workload involved in the data collection to generate the dashboard.*
36	Facility shame & culture of blame	NEST Interview 22 Line 100	*"I think there, of course, is danger around misinterpreting some of the data and what it means and what actions that might lead to. I think, also, I don't know. There could be issues around **facility shame**, I guess. Like hospital [administrators? - signal breaks up 0:32:09.3]. We do anonymise the facilities using IDs, but once you have data there and you see it, if there's something really bad, and you don't do anything about it there might be some... I don't know, I could see that there could be some issues, potentially, around that. Also, **a culture of blame.** I guess and, so far, the visits have been very constructive. This is something we also train our facilitators on. We're not there to assess. This is not an exam. No one should be pointing fingers, but I could also imagine that if something looks like it's the fault of one person that, yes, I don't know how that might play out."*
User Perspectives: Data Accuracy, Quality and Trust			
37	Data collection affecting trust	CIN Interview 17 Line 27-36	Respondent: The source document that the clerk use. I think that is the problem in the file it is written that one I can vote for it because it said during my ward round. But where this clerk gets the information is what I need. Although I have seen here we have…for the hospital it is just a general document. Moderator: Okay this is not necessarily for your…it's a general document. Respondent: Sometimes that is what I struggle with, I have tried to ask him where we do, we have talked about the discharge report on indications but [00:10:00] maybe also we would need somebody who clearly tries to tell everybody who comes in there because you know we are in rotations. The interns go, sometimes, the

			MOs go, the only steady person there is me probably and my colleague. Sometimes, somebody is on leave, like now I am on leave and the other person is there and the other person doesn’t have information. Even if a new paediatrician comes, they will not know if that is where you need this information to be picked from. Where the clerk picks the information from is key and important to us. It is something that also needs to be enhanced. Moderator: To follow that habit slightly over time, what are your experiences in terms of, you receive the report, you look at it and you know honestly, we have done discharges. Respondent: The last report I got from the clerk was… the main problem was on our nursing documentation because they look at the nursing report. That one of course, the nursing manager told me with the numbers he has chatting on that chat. Moderator: The monitoring chat? Respondent: The monitoring chart is a bit tricky for them, so they monitor yes, but the chatting is what is giving them a problem because sometimes they are only two with 45 children to make sure are fed, monitored with running everything, and sometimes it is tricky. Moderator: Two points I am picking from that conversation, one, is that there is a need for clarity on where, what the source documents are so that whatever is on the file is reflected on the report is an accurate reflection of what you’ve done in terms of documentation as well as no challenges with the rotation within the nursing team later in the monitoring chat, et cetera. It doesn’t mean that it’s not done. Respondent: That is the thing. It doesn’t mean that it is not done, it is done. Where it is written is the problem because the information is in the file. If I wanted it, I would… but now you see we also probably need to change towards the monitoring tool because it is what we use most of the time. Now considering that we also know that our hospitals are not equipped enough, especially with the nursing team, what else can we do?

38	Data collection affecting trust	Data clerk Interview 4 Line 100	M: I don't know what your thoughts are around. So there've been discussions around discrepancies as a result of where your source document, where you're sourcing the data from, I don't know what your experience have been? R: Sure, that is true. Like I just explained if he's sourcing his data from the register, I am sourcing mine from the patient file, you know this are two different sources of data, but created by same people you know, so and maybe the HRIO is actually sourcing hers maybe from the administration because at the end of the day, maybe what is from the wards is sent to the administration, so maybe she'll get hers from the nurse in charge or the in charges administration block. So you see all these people are getting their data from different sources and there are higher chances that one source could actually be missing something or having an incremental of something. So maybe if you asked me, probably just to cut that problem, maybe to just have one source of data so that you actually sit down and say its true. This is what we have. And we've agreed. So that issue of different sources I think is what is bring the discrepancies.
		Paediatrician Interview 18 Line 113-114	M: So, do you trust all the information that is given to you? R: So, there are instances where you are pretty sure you did a phenomenal job, you did a fantastic job but when you look at the details you are like wait, you are saying we did not do enough kangaroo? So we question and obviously we don't take everything that is here as gospel truth we have to check for ourselves and we also have to collate with our own information because we do record deaths and do all those things and if it doesn't tally we have to go and ask the person who collected the data "what exactly happened? Can you explain to us what happened?"
39	Mistrust due to failure to interpret	Nurse Interview 1 Line 44	R: There are challenges. I'll tell you there are challenges. If you were to talk to us as medical people, I think probably the issue of data analysis probably is not well covered. Although it is covered, yes, but using the technical now in terms of analysing, it might not be easy for somebody to interpret such a graph. But once you've got the opportunity to understand on how to interpret, then sharing to staffs

	rather than data quality		becomes easier. But then when it comes that now I want every staff to look at this data and discuss, in the very many forums we have gone, the QI and this NEST site facilities, that's the challenge you will find. I give you an example of this. You will give every participant represented in this dashboard, look at it. Yeah, they'll look, they'll look. But you ask them to interpret, it becomes a challenge. But it becomes easier when you tell them in the last three months by introducing them that this is a report which covers September to December, you had a total admission of this number, then you show them the total admissions in this graph. Then you tell them against this, you had a mortality rate of 19%. You give them a small calculation, what does 19 represent in this figure of 500? Even when you allow them to discuss and then you analyse for them, even the previous perceptions and answer could change. That's when they get to understand. Then, interestingly, they end up owning this document. They say, "Yes, that's ours." You know sometimes we have got that initial deny, "Is that really ours?" But when you try to break them down because you have interacted with these kinds of data and you understand is when they say, "Sure," and you end up even getting more information from here that is not here. They hacked your information from the ground. You find one or two looking at it and they say, "You remember, you remember, you remember," is when you start realising there is more than happens, but probably data does not come here because they also accept the gap, "There are some things we don't always record." But this one is for your consumption in this sense, there is data, they fear recording, because probably they'll get a backlash from the higher office. You'll be referred to conceal because it will give us a bad image. I remember one place we went, there was, there is another that quarterly summary report, there is somewhere a zero was written. This zero was written because from the standard reporting forms, there is no data that was collected. But in their own mind and in the ground, they know it was done, but because of what happened, they could not record it. I want to give a general, even if it happens to us, we report but we report with a lot of caution and security. For example, a maternal mortality, it will be known, you cannot hide but when you report you find you'll call for a meeting maybe 7:00 in the evening, for you to get right what led to this death? So that the moment you report, definitely even the national government you will find it in the morning at 8:00, what happened?

			But if you have not have done that, you put your things right so that when they come tomorrow, you have got something to explain, it is, so at times there are places you will find, this one should wait a little more, because it has implication because it is red tape, a maternal death is a red tape. Yes.
40	CIN data more accurate	HRIO Interview 16 Line 57	The CIN report is a bit more accurate than the reports we get. Reason being, the data clerk dealing with CIN would have to have the file, other than the other HRIO who could only record from the register where the nurses filled them. Normally, that health records officer would not note discrepancies in the report. But for the CIN data clerk, would be able to note and also the other checks and balances from the other side from KEMRI. So, if there are omissions or commissions, which they have not noted, it is picked out and they are able to communicate and acted upon in time. Other than ours…other than now he interrogations about the balances because we have interrogation now at the county level. Whereby maybe you reported you had [Inaudible 00:18:22] in neonatal death it does not balance with the number of children who are discharged plus the deaths should be equal to the number delivered. So, they only query when there is a discrepancy in that. Which might not actually be a discrepancy because if you had leftovers of last month, they would affect the report of this month, but the CIN is real time. That is the difference.
41	Not easy to realise data is wrong	Nurse Interview 1 Line 38	Even when you come now back to the lead you say, "We had a meeting, it was realised this is happening in our unit upon the data we submit." Because again, it is might not be easy for you to realise your own data that you're generating is wrong, but if somebody else looks at it he will tell you, here there's an issue. Yes. So that's how sometimes those now we call them sub county also targets come [unclear 27:36].
42	Valuable skills to identify errors	HRIO Interview 6 Line 152	Moderator: I don't know if there is something else you would like to tell me about the data use in bad report and basically calm to be the end of the question.

			Respondent: That was so nice I must say that CIN they are doing a good job and we commend them to continue with that good job. It has really improved the quality of data in our facility and also our officers even when they are transferred and go to another place, they are able to do the same what they have learn from this place. So, data is important and quality data is the best reconciliation. We advocate in this facility to ensure that every person is able to understand why data should be captured if everybody understands that you are able to”
43	Follow up on discrepancies	HRIO Interview 6 Line 30 -32	Moderator: I know you talked about you know that the expectation is sort of to ensure there is quality data, timely data as well, I don’t know, does these reports, the feedback reports help you in any way to achieve that your goal of ensuring that there is quality data? Respondent: Exactly, because CIN we have come from very far, earlier before, we were not having proper documentation and so our quality of data was very poor but when the CIN came in, we have seen a lot of improvement, we can be able to even see where we are not doing well and discussed together and improve. For me I can recommend the CIN group because they have really done well to us, especially in terms of data quality, I am able to stand still and be able to defect data that is from the CIN because it…even sometimes the data from CIN we checked together with our data and see if it is, it is a…going together. Moderator: How has that been, what are the experience been when you make compressions? Respondent: For CIN sometimes looks, genuine, because, actually they don’t miss any file, any documentation, it’s never missed but to us sometimes you can check around and leave a workload and at the end of the day, you can see why we don’t have 10 mortalities but from CIN, they are saying they were 12, you go back and check, so for me CIN is doing well and it has really help in the improvement of the quality data in this facility.

44	Data for advocacy	HRIO Interview 19 Line 176	M: You mentioned arguing to who? R: You see now arguing to the either administration or consultants because you see a death has occurred and it is your death and need to discuss, so when you discuss you say this death where did this patient come from? Was it a referral, yes. Was the referral prompt? Did the referral come in time or late? So that you give a report and once the county follows, "there are so many neo-natal death, why?
45	Data-driven interventions a task of clinicians	HRIO Interview 16 Line 65 - 67	"I think that is now the disconnect we have. Because you see the CIN report should inform the clinician and they should know from the CIN and from our reports. That is when now they should realize what interventions needs to be done [00:20:00] to be able to curb some of these challenges. But you see, without really a vocal person, it's not done. For example, if I say even in our reports like the audit of the [Inaudible 00:20:19] (or) ama the neonatal deaths. It is sort of really actually it is we record people who push other than come from the clinicians. Whereby after having such a report, it should actually come from the clinicians, not from me. Because from me, you see I only have the numbers. I know the kids are dying because of this, but they have the nitty gritty. They should be the ones who actually really require this to be able to hold up interventions for some of those diagnosis which are really killing the [Inaudible 00:21:07]."
User Perspectives: Experiences and Challenges with Dashboard Visualisations			
46	Interpretation difficulties	Nurse Interview 1 Line 44	*"There are challenges. I'll tell you there are challenges. If you were to talk <u>to us as medical people</u>, I think probably the issue of data analysis probably is not well covered. Although it is covered, yes, but using the technical now in terms of analysing, <u>it might not be easy for somebody to interpret such a graph.</u> But once you've got the <u>opportunity to understand on how to interpret, then sharing to staffs</u>*

			becomes easier. But then when it comes that now I want every staff to look at this data and discuss, in the very many forums we have gone, the QI and this NEST site facilities, that's the challenge you will find. I give you an example of this. You will give every participant represented in this dashboard, look at it. Yeah, they'll look, they'll look. But you ask them to interpret, it becomes a challenge. But it becomes easier when you tell them in the last three months by introducing them that this is a report which covers September to December, you had a total admission of this number, then you show them the total admissions in this graph. Then you tell them against this, you had a mortality rate of 19%. You give them a small calculation, what does 19 represent in this figure of 500? Even when you allow them to discuss and then you analyse for them, even the previous perceptions and answer could change. That's when they get to understand. Then, interestingly, they end up owning this document. They say, "Yes, that's ours." You know sometimes we have got that initial deny, "Is that really ours?" But when you try to break them down because you have interacted with these kinds of data and you understand is when they say, "Sure," and you end up even getting more information from here that is not here."
47	Interpretation: Need for guidance	Nurse Interview 2 Line 35	*Moderator Okay. In spaces where there have been presentation or interactions to staff or frontliners on this tool, I don't know whether there are areas you've picked that would be maybe say difficult to interpret, whether you've had challenges specifically with that cadre stage? Respondent 2: This one, once somebody has prepared it, he needs to give you an overview. Remember, not all of us are able to interpret graphs. We need to be taken through by the person who was made this.*
48	Interpretation: Need for training	Paediatrician Interview 13 Line 120	*"Besides stationery, training. I say training because sometimes I wonder do we interpret this tools the same way? Is the way I interpreted the same way as everybody across the county? I have this curious mind but just because of. Am I interpreting it the same way? Truth be told I don't know. Maybe I just keep interpreting it the wrong way and people assume I am the only one who is correct. Since I'm the senior most I should know. It is one big problem of medical care. People just assume because you are*

			the senior most doctor you know. … That is very far from the truth. That's why I say training as basic as it may seem maybe we should just have an induction hour or something. I was inducted to these tools by my colleague who have been here before assuming he was correct in his interpretation. Then along the way I also look and I'm like, "Okay this makes sense, this doesn't make sense." Occasionally we've heard during the CIN meetings, which again know not everyone attends. We've had sessions where we go through, I remember in Mombasa the CIN implementers asked a question and people laughed but it was a serious question. When he was asking, "Do we understand how to interpret, or we know like you guys are just feeding us with things? … I would put training as a need. As basic as it may seem, just having those 30 minutes to go through the report and help me synthesize it."
49	Simple design: More graphs vs one graph containing all information	Paediatrician Interview 18 Line 121	*"Now, it takes a lot of time to look at this and there are so many, like temperature okay, then the rest rate and for someone who doesn't have a very good eye, it takes a really long time to do that and as you keep looking at many of these are… some is off. …. "This is hard and it's not because it doesn't have good information but it's just that when there are many things overlapping [referring to lines in line chart] you get a little frustrated. You go and you are like what is this? I know having many graphs is chaotic but I would rather have two, maximum three that I can look at comfortably instead of having all the information in one, it took me a long time to look at that is probably the only thing I noted."*
50	Printing for presentations	HRIO Interview 6 Line 79	*We do not have a coloured, what (nini) painters, so sometimes we are not able to do our graphs well when we want to do a good presentation somewhere, we just print out white and black even the hospital computer doesn't have a coloured.*

Appendix 6.10. Quotes: System Integration and Contextual Influences, Section 3.3.

#	Theme	Respondent	Quote
User Perspectives: Feedback Delivery			
51	Timing of feedback	Nurse Interview 1 Line 83 - 90	*M1: Nice, that is okay. Other than the challenges that you've talked about of people not being familiar with data and analytics and the interpretations, are there any other challenges within where you work, your facility, to use of data or information tools here in the facility beyond now those technical interpretations of the staffs?* *R: I might not be worried more of the technical bits, but its challenge is having constant flow of that data feedback that says. Like now, CIN brings this data quarter.* *M1: Quarterly is like every?* *R: After every three months. That is, like now we have received the report of January, February March. So, April, May, June, we haven't received it. To me, I was talking even to one of the clerks here, is it possible, I was looking on a personal basis, being a deputy unit in charge, if I can be getting that report every month, a summary report that he sends via email, because it's my duty also to interpret how we are performing, I have a look at it, it can make me give feedback probably on a monthly basis when we have got our own meetings before we do it for such a report. Because somebody will tell you, you might not see your performance in one month, but on a small unit as the in charge, we need to know your progress every month. Even if you'll come to compile it after three months, but you have got some feedback you've already given to close the gaps so that we don't leave it for that quite long. If I realize this month temperatures taken percentage will increase, I can put immediate measures in the following month, for me at the end of three*

			months I'll have an inclusive report instead of us sitting, you say, "We are generating, we are generating," until the end of three months when you realise, percentage went down. *M1: Would it be a different thing from what, because I think CIN does it monthly, I don't know?* *R: I'm made to understand they do every month, but then it does not come to us. Once it goes to KEMRI, I think they analyse. But there is that daily generating data I think I'm aware of that the life, there is a name they give it, but it is somewhere there and probably from lab department. But then, either the in charge of the unit or the apartment, I don't think he gets that monthly summarised report.* *M1: So, even if it were monthly, you would still want a summarized format in terms of…?* *R: Yeah, because now when it comes to the issue of the other CRM in terms of now <u>you want to review your last three months, I am, it is here, that I can also improve performance that instead of me waiting until the end of the three months for us to do QI and see where we went wrong, some gaps can be dealt with within one month for the figures to improve.</u> Because again, data is very important in requesting change. There is that data that you can get in one week, it prompts you that if we continue this trend, we might be in the bad side by the end of the month. There are those easily adjustable challenges that you can do within a shorter time instead of waiting for us, or still waiting for maybe CIN to come and tell you this is how you've performed in the last month. This is good feedback, yes, but if you can be allowed to have that monthly picture, not for all staffs, maybe for the in charge, you see and protect, "Here, I don't want to wait until the end of the quarter and told that there was a mistake, I can rectify if that is possible." Then another issue is about erratic supply of data collection tools, that also affects data outcome. This time you have comprehensive care chart, the next one you're told there are no papers, you know, you end up, you want to create, if you are not able to create, you end up going to the cardex. The cardex, the data or summary might not be able to pick.*

User Perspectives - Resource constraints			
52	Adequate and continuous procurement	Nurse 1 Line 103-104	*M1: Anything else you want to ask that I have not asked?* *R: Not really, but I think we have discussed supplies and timely maintenance of equipment, we have when they break down, the consumables, that flow of consumable to be kind of continuous, because ins the performance the way we were discussing, sometimes it's affected by inadequate, not inadequate but inconsistent supplies. This week they are there, next week they are not there, the other time they are in excess, the other time they are not.*
53	Lack of stationary	Nurse Interview 11 Line 187	*"Lack of stationaries. that is number one. […] Number two is shortage of staff. You may find they might be there but time to chart is limited. So, you are one or two on shift, imagine doing nursing care, feeding the babies, giving medications charting those things or you find most of the time there are not even charted." M: "So, directly impacts on documentation?" Respondent: "Yes."*
54	Lack of stationary leading to loss of documentation	Nurse Interview 12 Line 119	*"Stationaries, like the whole of last year people have helped us. […] People now will prescribe on papers, but at times you do not get them. Then us guys…you don't have a comprehensive file. You have everything in papers, treatment sheet paper on its own, everything on its own. At times they get lost. They get lost, so maybe you need some information, continuous information of patient management."*
55	Referrals	Interview 18 Line 16	*"Say referrals for example yeah, for a number of referrals I have realized that we have been getting a certain number of referrals from a particular hospital. It may not give you exactly which hospital, but it will tell you the number of referrals are quite outnumbered compared to your own. That means there is a gap, for them to refer there is a gap in whatever facility that is referring and so it will help us say, okay identify what the gap is, what was the reason for the referral? "Pre-term", okay do they have a newborn unit or what was the*

			reason for that or let's say they did not have CPAP that is why they are referring to you and is it possible to empower those so that they provide quality care so that they don't refer patients?"
User Perspectives: Staff Turnover, Staff Shortage and Workload			
56	Staff turnover	Paediatrician Interview 13 Line 138	*"The report has been fine, we tend to get it by say, mid-month of the following month so that has been okay. We've been able to interpret it. <u>I think my biggest challenge has been this tough turnover and ability to implement change. By the time I get the previous month's report, the team has changed</u>. I have a new team. I grapple with that. Unfortunately, it's beyond CIN."*
57	Rotation of interns	Nurse Interview 17 Line 67	*"…my main problem is we need to have more integration between the nursing staff of course and the doctors. But at the end of three months, when I want to do a report and analyse all I have, most probably a medical intern has already rotated, who doesn't know what happened last month, and then maybe the nursing officer who has always been there who gets the report. Our turnover sometimes is a bit tricky because even if you are presenting data; sometimes you're presenting data for the previous group and another group."*
58	Extreme workload	Nurse Interview 5 Line 179	*<u>"Workload.</u> Like today we've come like we've reported four nurses against 74 babies out of the four you see now our in charge has purely been on administrative duties, so that leads three, now out of the three divide it amongst the 74 babies, so at what point will this person who is **[00:35:00]** taking care of these babies <u>leave them to sit down and discuss this data.</u> You find we don't get that time in fact, <u>you see tea, bread is here none of them has been able to come even to have break for tea cause its not possible there is a lot going on, so you find that the workload is so much that even getting a setup, a meeting to sit down and do a meeting it becomes a problem."</u>*

59	Prioritisation of tasks due to staff shortages and workload	Nurse Interview 2b Line 18-26.	*"Not really, everything is as we've talked the issues as I've mentioned workload, shortage of staffs. Everything revolves around there until you find somebody has not done it not intentional but situation. How the situation on the ground, things are different on the ground. You go now as per the situation. Am I going to leave a baby dying? Or am I going to continue resuscitation? To go and take vital observations? Which is more life threatening? Resuscitation I go, I deal with it. Observations I will take them later." …. "What I would say data is okay, it speaks for us. It advocates for us for more staffs or more resourceful things to be used in the care of our newborn babies if well utilised if well captured."*
60	Documentation challenge due to workload	Nurse Interview 5 Line 192	***"****Documentation there is a problem because of that same, workload you know that form is so, it is very much detailed and you find that one nurse is against 35 babies, 30 babies she might not be able to fill everything at least she'll be doing the most necessary, the most key, that is like OB's and the feeds they have to record every time that they, at least during the morning hours two time in a day, afternoon and at night twice. It's not as much as we would like it to be if we had a good, we had god ratios, yes….People are trying, they are really trying, they are really working hard despite all the challenges."*
61	Data to communicate and advocate for more staff	Nurse Interview 8 Line 159 – 160	***"****I would say that data can speak even to the big offices, I know my data also speaks to my boss, so even as she is praising nurses, she knows how much workload I have. It also communicates to the procurement, though they don't give you… like now I'm looking for diapers, though they might not give you what you actually require but they know that when if not given what you require maybe you are suffering…. I know it also communicate to the government even in terms of supplies. …because now if we interpret the data, the data will tell you in this stage here, people are not doing so nicely so they might need some support. Actually, you can even get support from the NEST team, they have really supported us in terms of capacity building. They have actually trained some nurses which actually comes back to us, and we improve our care. I can say it doesn't only communicate to one department, but it communicates to many departments."*

Chapter VII Discussion Appendices

Appendix 7.1. CP-FIT Hypotheses Checklist

For reflections on findings from quantitative and qualitative data I used the 42 CP-FIT hypotheses [55] as a checklist.

If not mentioned specifically otherwise, I am reporting findings in column 5 from both dashboards CIN and NEST as applying them to both tools. I differentiate between users and implementers, as well as between findings from the quantitative and qualitative data The colour-coded traffic light system is used to indicate the extent to which certain findings are addressed in the current dashboard design. It helps prioritise areas that need improvement or attention:

- ***Green:*** *These aspects are well-covered in the current design and do not require significant changes.*
- ***Yellow:*** *These areas are somewhat addressed but need further refinement or modification.*
- ***Red:*** *These elements are not adequately covered and require attention to improve the dashboard's effectiveness.*

CP-FIT Hypotheses: Feedback interventions are more effective when ...	Relevant feedback cycle process(es)	Key explanatory mechanism(s)	Paper reference from CP-FIT and thesis reference in [ ]	Findings from survey and interviews investigating the use of CIN & NEST dashboards
FEEDBACK VARIABLES				
Goal				
1. Importance: ... They focus on goals recipients believe to be meaningful and often do not happen in practice.	Acceptance, Intention	Compatibility, Credibility,		**Survey** Users perceive presented neonatal KPIs as important **Interviews** Implementers followed literature for selecting neonatal KPIs. Users highlight the importance of improving quality of care as one of their main aims in their role
2. Controllability: ... They focus on goals perceived to be	Acceptance, Intention	Actionability	Gort M et al (2013), CP-FIT 62, [218]	**Interviews** Improving KPIs can be challenging for users and out of their control

within the control of the recipients.				
3. Relevance: ... They focus on goals perceived as relevant to recipients' jobs.	Acceptance, Intention	Actionability, Compatibility, Relative advantage	Wilkinson EK at al (2000), CP-FIT 64, [256]	**Survey & Interviews** The KPIs from both dashboards are seen as relevant by users
Data collection and analysis method				
4. Conducted by recipients: ... They do not require the recipient to collect or analyse the clinical performance data. (Can also decrease 41. Cost)	Data collection and analysis	Complexity, Resource match	Cameron M et al (2007), CP-FIT 67, [219]	**Interviews** CIN developed and integrated a data system and employed data clerks. NEST "piggybacked" on CIN data system and did not create a system in parallel.
5. Automation: … They collect and analyse data automatically rather than manually.	Data collection and analysis	Complexity, Resource match	Palmer C et al (2012) CP-FIT 68, [220]	**Interviews** Data is analysed by implementers and then provide feedback to users.

6. **Accuracy:** …They use data believed by recipients to be a true representation of their clinical performance.	Acceptance	Credibility, Relative advantage	Payne V et al (2016), CP-FIT 50, [221]	**Interviews** There are trust issues related to both dashboards around data accuracy by users because of the underlying data quality issues and because of interpretability challenges. Implementers acknowledge the challenges around data accuracy.
7. **Exclusion:** … They allow recipients to exception report patients they feel are inappropriate to include in their performance measurement.	Acceptance	Actionability, Credibility	Chadwick et al (2016), CP-FIT 70, [222]	**Interviews** There is no exception report system, but users tend to discuss with colleagues if there seems something questionable and tend to report concerns to implementers.
Feedback display				
8. **Performance level**: ... They communicate recipients' current performance has room for improvement.	Intention, Behaviour	Actionability, Compatibility	Wilkinson EK at al (2000), CP-FIT 64, [256]	**Programme information** CIN included benchmarking to show there is room for improvement and a traffic light system. Otherwise, both tools do not show the average baseline directly to which users are supposed to 12improve their performance. However, users can see from both dashboards if they are performing poorly through specific outcome indicators and

				users also tend to look at these first, for example mortality rates.
9. **Patient lists:** ... They show the details of patients used to calculate the recipients' clinical performance.	Verification, Acceptance, Perception, Intention, Behaviour	Actionability, Complexity, Credibility	Payne V et al (2016), CP-FIT 50, [221]	**Programme information** No, there is not individual patient data visible.
10. **Specificity:** ... They report the performance of individual health professionals rather than their wider team or organisation.	Acceptance, Intention, Behaviour	Actionability	Nessim C (2012), CP-FIT 72, [257]	**Interviews** No, the feedback is on an organisations level. Implementers try to avoid the "blame game".
11. **Timeliness:** ... They use recent data to calculate recipients' current performance.	Acceptance, Intention, Behaviour	Actionability, Credibility	Payne V et al (2016), CP-FIT 50, [221]	**Interviews** Timeliness is something that comes up from implementers and users and is linked to the challenge of data collection systems in place. The aim is to have more timely data, maybe even real-time data, but currently the dashboards rely on retrospective data. Users mention that at the time they receive feedback the opportunity to change

				performance might be gone and some staff is not even working at the facility (frequent staff turnover).
12. Trend: ... They show recipients' current performance in relation to their past performance. (Can also increase 40. Observability)	Perception	Complexity, Relative advantage	Jeffs L et al (2014), CP-FIT 73, [225]	**Interviews** CIN presents current with past performance as part of the traffic light system. NEST does not compare with past performance. Users don't mention the past performance as incredibly beneficial.
13. Benchmarking: ... They compare recipients' current performance to that of other health professionals, organisations or regions.	Perception, Intention, Behaviour	Complexity, Social influence	Mannion R et al (1993), CP-FIT 74, [226]	**Interviews** CIN benchmarks with other facilities as part of the traffic light system. NEST does not compare with other facilities. Users appreciate comparisons with other facilities as it helps a) to know they are not alone in the challenge to improve performance, b) to learn from those who performed better
14. Prioritisation: ... They communicate the relative importance of feedback contents.	Perception	Complexity, Relative advantage	Yi S et al (2013), CP-FIT 55, [227]	**Interviews** The dashboards do not directly communicate prioritisation; however, this is discussed by users in team meetings based on the feedback. Both dashboards seem to trigger team meeting discussions to set priorities for performance/ quality of care improvement. User mostly look at

				mortalities first when they receive feedback from dashboards.
15. Usability: ... They employ user-friendly designs. (Can also increase 40. Observability)	Perception	Complexity	Guldberg, T et al (2010), CP-FIT 82, [258]	**Survey** **Interpretation success rates:** Both dashboards have some users who can interpret dashboard content, however, the survey showed that there are at least 50% of users who struggle to interpret the content and therefore will not be able to use the tools as intended. **Likert ratings Ease of Use & Usefulness**: When asked to rate how easy to use and useful the dashboards are, most respondents gave a high score indicating they are very easy to use and very useful.
				Interviews Implementers tried to design user-friendly dashboards. When users were asked during the interviews, they tend to confirm the Likert scale rating response, saying the dashboards are easy to use and useful, and graphs easy to interpret. However, later in the interviews mention the difficulties in interpreting the graphs.
Feedback delivery				

16. Function: ... They are perceived to support positive change rather than punish suboptimal performance.	Acceptance	Compatibility	Exworthy M et al (2003), CP-FIT 85, [259]	**Interviews** Both dashboards are not perceived as tools that punish bad performance. Although it can be frustrating to look at poor performance and high mortality rates, users seem to be motivated by the feedback to work on improvements.
17. Source knowledge and skill: ... They are delivered by a person or organisation perceived to have an appropriate level of knowledge or skill.	Acceptance	Credibility, Social influence	Creswell K et al (2012), CP-FIT 86, [260]	**Survey & Interviews** The CIN and NEST360 programmes are perceived as supportive and helpful. Participants mention in both, survey and interviews that they are grateful for the programmes and this thesis research.
18. Active delivery: ... They "push" feedback messages to recipients rather than requiring them to "pull". (Except if solely delivered face-to-	Interaction	Compatibility, Complexity		**Interviews** The feedback through both dashboards is delivered to the users, and as such can be seen as "push". However, people mention to require a print option, so there is some flexibility in terms of where the feedback is used. They also mention that they value the face-to-face interactions and team meetings. The dashboard feedback is seen like a supportive tool for these meetings.

face, which increases 41. Cost)				
19. Delivery to a group: ... They deliver feedback to groups of recipients.	Perception, Intention, Behaviour (by increasing 28. Teamwork)	Social influence	Hysong et al (2014), CP-FIT 98, [261]	**Interviews** The feedback is delivered to certain groups in the hospital that are engaged with neonatal care, to HRIOs and hospital management.
RECIPIENT VARIABLES				
Health professional characteristics				
20. Feedback attitude: They target health professionals with positive beliefs about feedback.	All	Compatibility, Relative Advantage	Wilkinson EK at al (2000), CP-FIT 64, [256]	**Interviews** Users have positive beliefs about feedback despite having some challenges with it. Implementers target people in senior positions, but not clear how they target users with positive attitudes. One user mentions worried about potential unintended consequences of the information which indicates there could be

21. Knowledge and skills in quality improvement: They target health professionals with greater capability in quality improvement.	Perception, Intention, Behaviour	Actionability, Complexity, Resource Match	Morrell C at al, CP-FIT 91, [262]	**Survey** Likert scale rating on confidence levels of users showed most of them to be confident to have the skills and knowledge to improve quality of care. However, the perception of **training needs** is linked to data and dashboard use.
				Interviews Respondents confirm the ratings from survey.
22. Knowledge and skills in clinical topic: They target health professionals with greater capability in the clinical topic under focus	Perception, Intention, Behaviour	Actionability, Resource match	Grant A et al (2014), CP-FIT 92, [263]	**Interviews** Implementers target people that are linked the newborn care.
Behavioural response				

23. Organisation-level and Patient-level behaviour: Health professionals undertake changes involving the wider health care system rather than just individual patients in response to feedback. (Can also increase 24. Resource).	Clinical Performance Improvement	Actionability	Powell A et al (2014), CP-FIT 95, [217]	**Interviews** Users undertake changes that target wider health systems improvement in addition to patient-level changes, for example working on the documentation and data quality, trying to advocate for the health workforce etc.
CONTEXT VARIABLES				
Organisation or team characteristics				
24. Resource: ... Organisations and teams have greater capacity to engage	All	Resource match	Hysong et al (2014), CP-FIT 98, [261]	**Interviews** Resource constraints were mentioned by users as an important negatively influencing factor on the use of data,

with them. (Can also increase 23. Organisation-level behaviour)				the use of the feedback given and their efforts to improve quality of care.
25. **Competing priorities:** ... Organisations and teams have minimal additional responsibilities.	All	Resource match, Compatibility	Ivers N et al (2014), CP-FIT 90, [264]	**Interviews** User mention competing priorities as part of their time constraints.
26. **Leadership support:** ... They are supported by senior managers. (Can also increase 23. Organisation-level behaviour)	All	Credibility, Resource match, Social Influence	Dixon-Woods M et al (2013), CP-FIT 87, [265]	**Interviews** Implementers mentioned that they strategically navigated the tool implementation with the facility leaders. However, users who are in a leadership position mention that they use a lot of their time to support junior staff but high turnover rates make it difficult to keep up with additional tasks
27. **Champions:** ... They are supported by individuals in the organisation dedicated to making it a success.	All	Credibility, Resource match, Social Influence	Palmer C et al (2012) CP-FIT 68, [220]	**Interviews** Implementers mentioned that they had champions who played an important part for successful implementation and uptake of dashboards.

28. **Teamwork:** ... They are implemented into organisations or teams whose members work together towards a common goal.	Perception, Intention, Behaviour	Actionability, Resource match, Social influence	Nessim C et al (2012), CP-FIT 72, [257]	**Interviews** Users mentioned the importance of discussing feedback in team meetings as improving quality of care is a team effort, also spanning across other hospital departments.
29. **Intra-organisational networks:** ... They are implemented into organisations or teams with strong internal communication channels.	Interaction, Perception, Intention, Behaviour	Actionability, Compatibility, Resource match, Social influence	Nouwens E et al (2015), CP-FIT 51, [266]	**Interviews** User mention to use additional meetings and WhatsApp group chats to reflect on feedback. It seems there are strong communication channels, however, it is not entirely clear how well these communication channels work.
30. **Extra-organisational networks:** ... They are implemented into organisations or teams that actively communicate with external bodies.	Perception, Intention, Behaviour	Actionability, Resource match	Creswell K et al (2012), CP-FIT 86, [260]	**Interviews** Users highlight the benefits a connection to other hospitals could have. CIN provides network meetings to discuss challenges where representatives from different partner hospitals come together. These meetings are highly valued by users.

31. **Workflow fit:** ... They fit alongside existing ways of working.	All	Compatibility, Complexity	Wilkinson EK at al (2000), CP-FIT 64, [256]	**Interviews** Users mention to be waiting for reports sometimes and the value of being able to print the reports, one reason being to take the feedback with them to read it after their shift. This is due to the high workload where taking time to read a long report doesn't fit in.
Patient population				
32. **Choice alignment:** ... They do not include patients who refuse aspects of care measured in the feedback in their calculations.	Acceptance, Intention	Actionability, Compatibility, Complexity	Powell A et al (2012), CP-FIT 105, [267]	**Interviews** CIN excludes extreme outliers from feedback as part of the data quality checks. However, it is not clear whether data from newborns for which parents refused case are measured in the feedback and/or included in calculations.
33. **Clinical appropriateness:** ... They do not include patients whose care cannot be safely optimised further.	Acceptance, Intention	Actionability, Compatibility, Complexity	Grando VT et al (2007), CP-FIT 148,	**Interviews** No clear results about this item.

Co-interventions				
34. **Peer discussion:** ... They encourage recipients discuss their feedback with peers. (Can also increase 28. Teamwork)	Perception, Intention	Complexity, Resource match, Social influence	Vachon B et al (2013), CP-FIT 89, [268]	**Interviews** Both dashboards seem to encourage team meetings and discussions around quality improvement strategies as mentioned by users. Some users wish to have more peer discussions with people from other facilities, see above as part of Benchmarking. Users also mentioned using WhatsApp to share striking findings from the dashboards. Integrating a communication tool to the digital dashboard version could be something to be discussed.
35. **Problem solving:** ... They help recipients identify and develop solutions to reasons for suboptimal performance (or support recipients to do so).	Perception	Actionability, Compatibility, Complexity, Resource match	Ivers N et al (2014), CP-FIT 90, [264]	**Interviews** I tried to investigate this topic, but finding straightforward answers was difficult. When I asked users what they do with the feedback, the standard response was often, "Well, then I know." This left me questioning, "Then you know what? How can you use that to improve quality?" While the dashboards do provide feedback and help users understand their performance, they don't include specific tasks to take immediate action. Users mentioned that they often proactively followed up to understand where the data came from and identify potential errors they could address

				when a) the documentation KPI showed poor performance, or b) the clinical KPIs showed very poor results and data sources were questioned. When clinical KPIs showed poor performance, participant usually mentioned being overworked and understaffed. The tools were then mentioned as help to advocate for more staff and resources.
36. **Action planning**: ... They provide solutions to suboptimal performance (or support recipients to do so).	Intention, Behaviour	Actionability, Complexity, Resource match	Gort M et al (2013), CP-FIT 62, [218]	**Interviews** For further action planning, users mentioned that meetings and discussions with others are necessary to find strategies on addressing performance and improve quality of care. The dashboards are not providing solutions but serve as an assisting tool for these meetings and development of strategies. Providing some sort of action plan feature to the digital dashboard could be something to be discussed.
37. **External change agents**: ... They provide additional staff to explicitly	All	Resource match	Tierney S et al (2014), CP-FIT 94, [269]	**Interviews** Both, the CIN and NEST team provide support, for example through the QI visits or in-person meetings. However, the users mentioned there could be more support in terms of explaining the graphs. The senior nurses who already have a high workload are the main

support its implementation.				people who train junior staff, and since they have a high turnover rate, this seems to be a repetitive effort.
Implementation process				
38. **Adaptability:** ... They are tailored to the specific needs of the health care organisation and its staff. (Can also increase 31. Workflow fit)	All	Compatibility, Complexity	Groene O et al (2008), WHO Assessment tool, CP-FIT 69, [270]	**Interviews** Implementers mention that they carefully assessed the organisation before implementing the tools. NEST mentioned they conducted health facility assessments for each partner hospital. However, CIN implementers mentioned that the dashboards have not been tested in detail like this study does and appreciate this study will provide more details about the implementation processes too. The adaptability of the dashboard design is something that came up from users since needs may change and depend on a changing environment, e.g., staff turnover, strikes etc.

39. **Training and support**: ... They provide training and support regarding feedback (not the clinical topic under scrutiny).	Perception, Intention, Behaviour (by increasing 21. Knowledge and skills in quality improvement)	Actionability, Resource match	Morrell C at al, CP-FIT 91, [262]	**Survey** The Likert scale rating that asked participants to indicate their training needs showed that people seem to perceive to have more training needs and there was a negative correlation with the graph interpretation success rate. **Interviews** Implementers mention that they provided training and explanations on how to use the tools. However, users say they could have more training and wish that implementers help more.
40. **Observability:** ... They demonstrate their potential benefits to recipients.	All	Relative advantage	Turner S (2014), CP-FIT 88, [271]	**Survey** The Likert scale rating on usefulness showed high rankings, however did not correlate with the ability to interpret graphs from the dashboard. **Interviews** The users confirm that both tools are useful and they benefit from having feedback, although there can be some improvements on the tools.

41. **Cost:** ... They are considered inexpensive to deploy in terms of time, human or financial resources.	All	Resource match	Cameron M et al (2007), CP-FIT 67, [219]	**Interviews** Financial costs linked to the dashboards were not mentioned by any participant. Financial losses were mentioned by one user after I stopped the recording link to other health information tools that were implemented, such as for example Electronic Patient Records, that were not interoperable and therefore useless. Human costs were indirectly mentioned as additional tasks that come with data collection and training junior staff.
42. **Ownership:** ... Recipients feel they “own” it, rather than it has been imposed on them.	All	Compatibility	Mcdonald R et al (2009), CP-FIT 149, [272]	**Interviews** Implementers mentioned that they observed user “own” the tool and that was the intention too. However, some users mentioned that they think the implementers are the owners and therefore have the responsibility to provide more explanations and work on updates regularly.

Appendix 7.2. Conference Presentations & Research Community Feedback

Date	Forum	Location
05/2022	**Kenya Paediatric Association (KPA), Annual Scientific Conference**; Informal conversations with paediatricians and collaborators CIN and NEST	Mombasa, Kenya
10/2022	**Centre for Global Health Research (CGHR) Annual Conference**, Nuffield Department of Medicine; Oral (rapid) presentation	Oxford, UK
11/2022	**7th Global Symposium of Health Systems Research,** Health Systems Global (HSG), Poster presentation	Bogota, Colombia
09/2023	**Annual Conference on Global Health**, Richard Doll Society, Green Templeton College; Oral and poster presentation	Oxford, UK
11/2023	**Oxford Digital Festival**, Jesus College; Poster presentation	Oxford, UK
03/2024	**Global Child Health Symposium,** Green Templeton College; Poster presentation & awarded best research poster	Oxford, UK
04/2024	**BMJ International Forum on Quality and Safety in Healthcare;** Poster presentation	London, UK
06/2024	**Design4Health Conference**; Accepted for oral presentation, but due to lack of funding not attended	Sheffield, UK
11/2024	**8th Global Symposium of Health Systems Research,** Health Systems Global (HSG); Oral presentation accepted	Nagasaki, Japan
Regularly	**Regular work in progress (WiP) presentations;** Health Systems Collaborative (my research group) Learning Sessions & Peter Medawar Building Seminars	Oxford, UK

X. Bibliography

[1] C. Hagel, Y. Zhao, C. Paton, S. Molyneux, and M. English, “Exploring the development and use of dashboards in neonatal care: A scoping review protocol,” Sep. 2024, doi: 10.17605/OSF.IO/EVQZP.

[2] W. T. Gude *et al.*, “Clinical performance comparators in audit and feedback: A review of theory and evidence,” *Implementation Science*, vol. 14, no. 1, pp. 1–14, 2019, doi: 10.1186/s13012-019-0887-1.

[3] G. La Grouw, *Effective Dashboard Design - Design secrets to getting more value from performance dashboards*. 2012.

[4] “Defining key performance indicators (KPIs) - IBM Documentation.” Accessed: Oct. 12, 2024. [Online]. Available: https://www.ibm.com/docs/en/business-monitor/8.5.6?topic=models-defining-kpis

[5] J. Mant, “Process versus outcome indicators in the assessment of quality of health care,” 2001.

[6] The World Health Organization, “Quality of care,” Fact Sheet. Accessed: Sep. 14, 2024. [Online]. Available: https://www.who.int/health-topics/quality-of-care#tab=tab_1

[7] J. E. Lawn, Z. A. Bhutta, C. Ezeaka, and O. D. Saugstad, “Ending Preventable Neonatal Deaths: Multicountry Evidence to Inform Accelerated Progress to the Sustainable Development Goal by 2030,” *Neonatology Global Neonatal Research Neonatology*, vol. 120, pp. 491–499, 2023, doi: 10.1159/000530496.

[8] “Health - United Nations Sustainable Development.” Accessed: Oct. 02, 2024. [Online]. Available: https://www.un.org/sustainabledevelopment/health/

[9] L. Hug *et al.*, “Levels & Trends in Child Mortality - Report 2023.”

[10] “Newborn mortality.” Accessed: Oct. 02, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/newborn-mortality

[11] “Every Newborn Action Plan (ENAP) and Ending Preventable Maternal Mortality (EPMM) Dashboard .” Accessed: Oct. 02, 2024. [Online]. Available: https://platform.who.int/data/maternal-newborn-child-adolescent-ageing/ENAP-EPMM-dashboard

[12] World Health Organization (WHO) & United Nations Children’s Fund (UNICEF), “EVERY NEWBORN ACTION PLAN,” Country Progress Tracking Report. Accessed: Jan. 12, 2019. [Online]. Available: https://www.healthynewbornnetwork.org/hnn-content/uploads/ENAP-country-submissions.pdf

[13] World Health Organization, “Monitoring the building blocks of health systems: A handbook of indicators and their measurement strategies,” 2010. [Online]. Available: https://www.who.int/healthinfo/systems/WHO_MBHSS_2010_full_web.pdf

[14] “Monitoring health for the SDGs, Sustainable Development Goals.”

[15] C. Hagel, C. Paton, G. Mbevi, and M. English, “Data for tracking SDGs: Challenges in capturing neonatal data from hospitals in Kenya,” *BMJ Glob Health*, vol. 5, no. 3, pp. 1–10, 2020, doi: 10.1136/bmjgh-2019-002108.

[16] S. Gachau *et al.*, “Does audit and feedback improve the adoption of recommended practices? Evidence from a longitudinal observational study of an emerging clinical network in Kenya,” *BMJ Glob Health*, vol. 2, no. 4, 2017, doi: 10.1136/bmjgh-2017-000468.

[17] F. Odekunle, R. Odekunle, and S. Shankar, “Why sub-Saharan Africa lags in electronic health record adoption and possible strategies to increase its adoption in this region,” *Int J Health Sci (Qassim)*, vol. Vol. 11, Issue 4, Oct. 2017.

[18] D. Wickremasinghe, I. E. Hashmi, J. Schellenberg, and B. I. Avan, “District decision-making for health in low-income settings: A systematic literature review,” *Health Policy Plan*, vol. 31, pp. ii12–ii24, Sep. 2016, doi: 10.1093/heapol/czv124.

[19] “DHIS2,” In action. [Online]. Available: https://www.dhis2.org/inaction

[20] T. Tuti *et al.*, “Improving in-patient neonatal data quality as a pre-requisite for monitoring and improving quality of care at scale: A multisite retrospective cohort study in Kenya,”

PLOS Global Public Health, vol. 2, no. 10, Oct. 2022, doi: 10.1371/journal.pgph.0000673.

[21] J. H. Cross *et al.*, “Neonatal inpatient dataset for small and sick newborn care in low- and middle-income countries: systematic development and multi-country operationalisation with NEST360,” *BMC Pediatr*, vol. 23, Nov. 2023, doi: 10.1186/s12887-023-04341-2.

[22] NEST360 Alliance, “NEST360 Website - What we do.” Accessed: Sep. 10, 2024. [Online]. Available: https://nest360.org/

[23] T. Tuti *et al.*, “Improving documentation of clinical care within a clinical information network: An essential initial step in efforts to understand and improve care in Kenyan hospitals,” *BMJ Glob Health*, vol. 1, no. 1, pp. 1–7, 2016, doi: 10.1136/bmjgh-2016-000028.

[24] T. Tuti *et al.*, “A systematic review of electronic audit and feedback: Intervention effectiveness and use of behaviour change theory,” *Implementation Science*, vol. 12, no. 1, 2017, doi: 10.1186/s13012-017-0590-z.

[25] S. Few, *Information Dashboard Design - Displaying data for at-a-glance monitoring*, Second Edi. Burlingame, California, USA: Analytics Press, 2013.

[26] D. Norman, *User Centered System Design: New Perspectives on Human Computer Interaction*. 1986.

[27] United Nations Sustainable Development Solutions Network (SDSN), “Indicators and a Monitoring Framework - Launching a data revolution for the Sustainable Development Goals.” Accessed: Nov. 22, 2020. [Online]. Available: https://indicators.report/targets/

[28] World Health Organization (WHO), “World Health Data Platform/ Data collection tools/ SCORE.” [Online]. Available: https://www.who.int/data/data-collection-tools/score

[29] World Health Organisation (WHO), “Making Every Baby Count Audit and review of stillbirths and neonatal deaths.,” *WHO Library Cataloguing-in-Publication Data*, p. 144, 2016, [Online]. Available: http://www.who.int/about/

[30] "Neonate: MedlinePlus Medical Encyclopedia." Accessed: Oct. 02, 2024. [Online]. Available: https://medlineplus.gov/ency/article/002271.htm#

[31] G. Irimu *et al.*, "Neonatal mortality in Kenyan hospitals: A multisite, retrospective, cohort study," *BMJ Case Rep*, vol. 6, no. 5, May 2021, doi: 10.1136/bmjgh-2020-004475.

[32] K. National Bureau of Statistics Nairobi, "Kenya Demographic and Health Survey 2022 Key Indicators Report," 2023. [Online]. Available: www.DHSprogram.com.

[33] The World Bank, "Mortality rate, neonatal (per 1,000 live births) - Kenya | Data." Accessed: Sep. 14, 2024. [Online]. Available: https://data.worldbank.org/indicator/SH.DYN.NMRT?locations=KE

[34] M. E. Kruk *et al.*, "High-quality health systems in the Sustainable Development Goals era: time for a revolution," Nov. 01, 2018, *Elsevier Ltd*. doi: 10.1016/S2214-109X(18)30386-3.

[35] S. M. Campbell, M. O. Roland, and S. A. Buetow, "Defining quality of care," 2000. [Online]. Available: www.elsevier.com/locate/socscimed

[36] World Health Organisation, "Standards for Improving Quality of Maternal and Newborn Care in Health Facilities," *Standards for improving quality of maternal and newborn care in health facilities*, p. 73, 2016, doi: 978 92 4 151121 6.

[37] E. K. Diego, D. E. Y. Ehret, K. C. Ashish, and C. L. Bose, "Quality Indicators to Evaluate Essential Newborn Care in Low- and Middle-Income Countries," *Pediatrics*, vol. 152, no. 3, Sep. 2023, doi: 10.1542/peds.2023-061527.

[38] World Health Organization, "The Network for Improving Quality of Care for Maternal, Newborn, and Child Health - EVALUATION, IMPLEMENTATION AND PROGRESS 2017 - 2020 REPORT," 2021.

[39] The World Health Organization, "Quality, Equity, Dignity: A Network for Improving Quality of Care for Maternal, Newborn and Child Health," 2019.

[40] World Health Organization (WHO), "Every Newborn Action Plan Coverage Target to 2025 Online Global Consultation." [Online]. Available: https://www.who.int/health-

topics/newborn-health/every-newborn-action-plan-coverage-target-to-2025-online-global-consultation

[41] R. Lundin, I. Mariani, K. Peven, L. T. Day, and M. Lazzerini, “Quality of routine health facility data used for newborn indicators in lowand middleincome countries: A systematic review,” *J Glob Health*, vol. 12, 2022, doi: 10.7189/jogh.12.04019.

[42] K. J. Muthoka, “Why are Neonates Dying? Socioeconomic and Proximate Determinants of Neonatal Mortality among Stable Low-Birth-Weight (LBW) Infants in Kenya,” Simon Karanja School of Public Health, 2017. [Online]. Available: www.iiste.org

[43] A. B. Patel *et al.*, “The Global Network Socioeconomic Status Index as a predictor of stillbirths, perinatal mortality, and neonatal mortality in rural communities in low and lower middle income country sites of the Global Network for Women’s and Children’s Health Research,” *PLoS One*, vol. 17, no. 8 August, Aug. 2022, doi: 10.1371/journal.pone.0272712.

[44] The World Health Organization, “Everybody’s business: strengthening health systems to improve health outcomes: WHO’s framework for action,” World Health Organization, 2007.

[45] S. P. Zodpey and H. N. Negandhi, “Improving the quality and use of routine health data for decision-making,” *Indian J Public Health*, vol. 60, no. 1, pp. 1–3, Jan. 2016, doi: 10.4103/0019-557X.177248.

[46] M. Prince, “‘If you can’t measure it, you can’t manage it’ – essential truth, or costly myth?,” Feb. 01, 2018, *Blackwell Publishing Ltd*. doi: 10.1002/wps.20477.

[47] A.-S. Lu, Yonglong; Nakicenovic, Nebojsa; Visbeck, Martin ; Stevance, “Five Priorities for the UN SDGs,” *Nature*, vol. 520, no. April 2015, pp. 432–433, 2015.

[48] D. W. Bates *et al.*, “Using information systems to measure and improve quality,” *Int J Med Inform*, vol. 53, no. 2–3, pp. 115–124, 1999, doi: 10.1016/S1386-5056(98)00152-X.

[49] World Health Organisation (WHO), “Mother and Newborn Information for Tracking Outcomes and Results (MONITOR) technical advisory group,” Maternal, newborn, child

and adolescent health. Epidemiology, monitoring and evaluation. Accessed: Nov. 23, 2020. [Online]. Available: https://www.who.int/maternal_child_adolescent/epidemiology/monitor/en/

[50] H. L. Colquhoun *et al.*, "Advancing the literature on designing audit and feedback interventions: Identifying theory-informed hypotheses," *Implementation Science*, vol. 12, no. 1, pp. 1–10, 2017, doi: 10.1186/s13012-017-0646-0.

[51] Z. Landis-Lewis *et al.*, "Designing clinical practice feedback reports: Three steps illustrated in Veterans Health Affairs long-term care facilities and programs," *Implementation Science*, vol. 15, no. 1, pp. 1–13, 2020, doi: 10.1186/s13012-019-0950-y.

[52] H. Atherton, J. Fleming, V. Williams, and J. Powell, "Online patient feedback: a cross-sectional survey of the attitudes and experiences of United Kingdom health care professionals," *J Health Serv Res Policy*, vol. 24, no. 4, pp. 235–244, 2019, doi: 10.1177/1355819619844540.

[53] S. J. Hysong, C. R. Teal, M. J. Khan, and P. Haidet, "Improving quality of care through improved audit and feedback," *Implementation Science*, vol. 7, no. 1, p. 1, 2012, doi: 10.1186/1748-5908-7-45.

[54] Z. Landis-Lewis, J. C. Brehaut, H. Hochheiser, G. P. Douglas, and R. S. Jacobson, "Computer-supported feedback message tailoring: Theory-informed adaptation of clinical audit and feedback for learning and behavior change," *Implementation Science*, vol. 10, no. 1, pp. 1–12, 2015, doi: 10.1186/s13012-014-0203-z.

[55] B. Brown *et al.*, "Clinical Performance Feedback Intervention Theory (CP-FIT): A new theory for designing, implementing, and evaluating feedback in health care based on a systematic review and meta-synthesis of qualitative research," *Implementation Science*, vol. 14, no. 1, pp. 1–25, 2019, doi: 10.1186/s13012-019-0883-5.

[56] S. Michie, M. M. van Stralen, and R. West, "The behaviour change wheel: A new method for characterising and designing behaviour change interventions," *Implementation Science*, vol. 6, no. 1, Apr. 2011, doi: 10.1186/1748-5908-6-42.

[57] W. T. Gude and N. Peek, “Control Theory to Design and Evaluate Audit and Feedback Interventions,” *Stud Health Technol Inform*, vol. 263, pp. 159–170, 2019, doi: 10.3233/SHTI190120.

[58] J. Y. Tsang, N. Peek, I. Buchan, S. N. Van Der Veer, and B. Brown, “Systematic review and narrative synthesis of computerized audit and feedback systems in healthcare,” Jun. 01, 2022, *Oxford University Press*. doi: 10.1093/jamia/ocac031.

[59] W. H. Organization, *Standards for improving the quality of care for small and sick newborns in health facilities*. Geneva: World Health Organization, 2020. [Online]. Available: https://iris.who.int/handle/10665/334126

[60] H. Gannon *et al.*, “Electronic application to improve management of infections in low-income neonatal units: Pilot implementation of the NeoTree beta app in a public sector hospital in Zimbabwe,” *BMJ Open Qual*, vol. 10, no. 1, Jan. 2021, doi: 10.1136/bmjoq-2020-001043.

[61] T. Lovell, “The rise of the dashboard | Healthcare IT News.” Accessed: Sep. 14, 2024. [Online]. Available: https://www.healthcareitnews.com/news/emea/rise-dashboard

[62] “dashboard | Etymology of dashboard by etymonline.” Accessed: Oct. 04, 2024. [Online]. Available: https://www.etymonline.com/word/dashboard

[63] M. Batty, “A perspective on city dashboards,” 2015, *Routledge*. doi: 10.1080/21681376.2014.987540.

[64] D. P. Kaplan, Robert S.; Norton, “Using the Balanced Scorecard as a Strategic Managment System,” *Harv Bus Rev*, vol. 88, no. 2, p. 534, 2007, doi: 10.1016/j.fertnstert.2007.05.031.

[65] O. M. Yigitbasioglu and O. Velcu, “A review of dashboards in performance management: Implications for design and research,” *International Journal of Accounting Information Systems*, vol. 13, no. 1, pp. 41–59, 2012, doi: 10.1016/j.accinf.2011.08.002.

[66] A. Pace and S. C. Buttigieg, “Can hospital dashboards provide visibility of information from bedside to board? A case study approach,” *Journal of Health, Organisation and Management*, vol. 31, no. 2, pp. 142–161, 2017, doi: 10.1108/JHOM-11-2016-0229.

[67] M. Ghazisaeidi, R. Safdari, M. Torabi, M. Mirzaee, J. Farzi, and A. Goodini, “Development of performance dashboards in healthcare sector: Key practical issues,” *Acta Informatica Medica*, vol. 23, no. 5, pp. 317–321, 2015, doi: 10.5455/aim.2015.23.317-321.

[68] W. K. Van Deen *et al.*, “Involving end-users in the design of an audit and feedback intervention in the emergency department setting - A mixed methods study,” *BMC Health Serv Res*, vol. 19, no. 1, pp. 1–13, 2019, doi: 10.1186/s12913-019-4084-3.

[69] D. Dowding, J. Merrill, and D. Russell, “Using Feedback Intervention Theory to Guide Clinical Dashboard Design,” *AMIA Annu Symp Proc*, vol. 2018, pp. 395–403, 2018.

[70] A. Grant *et al.*, “Integrating feedback from a clinical data warehouse into practice organisation,” *Int J Med Inform*, vol. 75, no. 3-4 SPEC. ISS., pp. 232–239, 2006, doi: 10.1016/j.ijmedinf.2005.07.037.

[71] D. Dowding *et al.*, “Dashboards for improving patient care: Review of the literature,” *Int J Med Inform*, vol. 84, no. 2, pp. 87–100, 2015, doi: 10.1016/j.ijmedinf.2014.10.001.

[72] J. S. Ancker, Y. Senathirajah, R. Kukafka, and J. B. Starren, “Design Features of Graphs in Health Risk Communication: A Systematic Review,” *Journal of the American Medical Informatics Association*, vol. 13, no. 6, pp. 608–618, 2006, doi: 10.1197/jamia.M2115.

[73] H. van Elten, S. Sulz, R. Wehrens, and E. van Raaij, “Big Data Health Care Innovations: Performance Dashboarding as a Process of Collective Sensemaking,” *JMIR Res Protoc*, vol. 9, no. 10, Oct. 2020, doi: 10.2196/16779.

[74] C. Knaflic Nussbaumer, *Storytelling with data - A data visualization guide for business professionals*. New Jersey, USA: Wiley, 2015.

[75] R. A. Flinchum and R. O. Meyer, *Henry Dreyfuss: Designing for People*. State University of New York Press, 2022. doi: 10.2307/jj.18252727.

[76] D. A. Norman, “The design of everyday things,” 2011, *Tantor Media, Old Saybrook, CT*.

[77] J. Canny, “The Future of Human-Computer Interaction,” 2006. [Online]. Available: www.acmqueue.com

[78] G. M. Olson and J. S. Olson, “Human-Computer Interaction: Psychological Aspects of the Human Use of Computing,” *Annu Rev Psychol*, vol. 54, no. 1, pp. 491–516, Feb. 2003, doi: 10.1146/annurev.psych.54.101601.145044.

[79] “What is Human-Computer Interaction (HCI)? | IxDF.” Accessed: Oct. 04, 2024. [Online]. Available: https://www.interaction-design.org/literature/topics/human-computer-interaction

[80] “ISO 9241-210:2019 - Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems.” Accessed: Sep. 14, 2024. [Online]. Available: https://www.iso.org/standard/77520.html

[81] H. A. Kheder, “HUMAN-COMPUTER INTERACTION: ENHANCING USER EXPERIENCE IN INTERACTIVE SYSTEMS,” *Kufa Journal of Engineering*, vol. 14, no. 4, pp. 23–41, Oct. 2023, doi: 10.30572/2018/KJE/140403.

[82] P. Rajesh, M. Selvadurai, V. Saranya Selvamani, and P. Chandrasekar, *Fundamentals of UX/UI (An Approach to Design Principles)*. Jupiter Publications Consortium, 2022. doi: 10.47715/JPC.B.87.2022.9789391303389.

[83] Interaction Design Foundation, “What is User Experience (UX) Design? | IxDF.” Accessed: Sep. 15, 2024. [Online]. Available: https://www.interaction-design.org/literature/topics/ux-design

[84] Interaction Design Foundation, “What is User Interface (UI) Design? — updated 2024 | IxDF.” Accessed: Sep. 15, 2024. [Online]. Available: https://www.interaction-design.org/literature/topics/ui-design

[85] B. Shneiderman, *Designing the user interface : strategies for effective human-computer interaction*, Sixth edition. Boston: Pearson, 2017.

[86] J. Nielsen, *Usability engineering*, 1st edition. in Interactive Technologies. Cambridge, Mass: AP Professional, 1993.

[87] Nielsen Norman (NN) Group, “User Experience 2008 Amsterdam Conference.” Accessed: Oct. 05, 2024. [Online]. Available: http://www.neospot.se/usability-vs-user-experience/

[88] C. Rohrer, "When to Use Which User-Experience Research Methods." Accessed: Oct. 08, 2024. [Online]. Available: https://www.nngroup.com/articles/which-ux-research-methods/

[89] S. Gibbons, "Empathy Mapping: The First Step in Design Thinking," Nielsen Norman Group. Accessed: Sep. 16, 2024. [Online]. Available: https://www.nngroup.com/articles/empathy-mapping/

[90] R. F. Dam and T. Y. Siang, "Empathy Map – Why and How to Use It | IxDF," Interaction Design Foundation. Accessed: Sep. 16, 2024. [Online]. Available: https://www.interaction-design.org/literature/article/empathy-map-why-and-how-to-use-it

[91] E. de Lera, "Emotion-Centered-Design (ECD) New Approach for Designing Interactions that Matter," 2015, pp. 406–416. doi: 10.1007/978-3-319-20886-2_38.

[92] "Design Thinking 101." Accessed: Oct. 07, 2024. [Online]. Available: https://www.nngroup.com/articles/design-thinking/

[93] S. Few, *Now you can see it - Simple visualization techniques for quantitative analysis*. Burlingame, California, USA: Analytics Press, 2009.

[94] S. K. Card, *Readings in Information Visualization: Using Vision to Think*. 1999.

[95] Stephen. Few, *Show me the numbers - Designing tables and graphs to enlighten*. Burlingame, California, USA: Analytics Press, 2012.

[96] "The Extreme Presentation(tm) Method." Accessed: Sep. 22, 2024. [Online]. Available: https://extremepresentation.typepad.com/

[97] "Visual Vocabulary." Accessed: Sep. 22, 2024. [Online]. Available: http://ft-interactive.github.io/visual-vocabulary/

[98] E. R. Tufte, *The Visual Display of Quantitative Information*, 2nd ed. 2018.

[99] A. Smith, "Visual vocabulary - Designing with data." [Online]. Available: https://ft-interactive.github.io/visual-vocabulary/

[100] Stephen. Few, *Show me the numbers - Designing tables and graphs to enlighten*. Burlingame, California, USA: Analytics Press, 2012.

[101] C. LeRouge *et al.*, "Using Heuristic Evaluation to Enhance the Visual Display of a Provider Dashboard for Patient-Reported Outcomes," *eGEMs (Generating Evidence & Methods to improve patient outcomes)*, vol. 5, no. 2, p. 6, 2017, doi: 10.13063/2327-9214.1283.

[102] W. S. Cleveland and R. McGill, "Graphical perception: Theory, experimentation, and application to the development of graphical methods," *J Am Stat Assoc*, vol. 79, no. 387, pp. 531–554, 1984, doi: 10.1080/01621459.1984.10478080.

[103] T. Zuk, L. Schlesier, P. Neumann, M. S. Hancock, and S. Carpendale, "Heuristics for information visualization evaluation," *Proceedings of BELIV'06: BEyond time and errors - novel EvaLuation methods for Information Visualization. A workshop of the AVI 2006 International Working Conference*, pp. 1–6, 2006, doi: 10.1145/1168149.1168162.

[104] M. Hegarty, "The cognitive science of visual-spatial displays: Implications for design," *Top Cogn Sci*, vol. 3, no. 3, pp. 446–474, Jul. 2011, doi: 10.1111/j.1756-8765.2011.01150.x.

[105] T. Wagemans, J, Feldman J Gepshtein SFoundations, "A Century of Gestalt Psychology," vol. 138, no. 6, pp. 1218–1252, 2013, doi: 10.1037/a0029334.A.

[106] D. G. Dorner, *Information Needs Analysis - Principles and practice in information organizations*. London (UK), 2015.

[107] M. English, J. Schellenberg, and J. Todd, "Policy & practice Assessing health system interventions: key points when considering the value of randomization," *Bull World Health Organ*, vol. 89, pp. 907–912, 2011, doi: 10.2471/BLT.11.089524.

[108] J. Y. Tsang, N. Peek, I. Buchan, S. N. Van Der Veer, and B. Brown, "Systematic review and narrative synthesis of computerized audit and feedback systems in healthcare," Jun. 01, 2022, *Oxford University Press*. doi: 10.1093/jamia/ocac031.

[109] S. Almasi, R. Rabiei, H. Moghaddasi, and M. Vahidi-Asl, "Emergency Department Quality Dashboard; a Systematic Review of Performance Indicators, Functionalities, and Challenges," *Arch Acad Emerg Med*, vol. 9, no. 1, pp. 1–11, 2021, doi: 10.22037/aaem.v9i1.1230.

[110] R. Rabiei and S. Almasi, “Requirements and challenges of hospital dashboards: a systematic literature review,” *BMC Med Inform Decis Mak*, vol. 22, no. 1, Dec. 2022, doi: 10.1186/s12911-022-02037-8.

[111] J. McCool *et al.*, “Factors influencing the sustainability of digital health interventions in low-resource settings: Lessons from five countries,” *J Glob Health*, vol. 10, no. 2, Dec. 2020, doi: 10.7189/JOGH.10.020396.

[112] A. B. Labrique *et al.*, “Best practices in scaling digital health in low and middle income countries,” Nov. 03, 2018, *BioMed Central Ltd*. doi: 10.1186/s12992-018-0424-z.

[113] S. S. Kaboré *et al.*, “Barriers and facilitators for the sustainability of digital health interventions in low and middle-income countries: A systematic review,” Nov. 28, 2022, *Frontiers Media S.A.* doi: 10.3389/fdgth.2022.1014375.

[114] N. M. Ivers and J. Barrett, “Using report cards and dashboards to drive quality improvement: Lessons learnt and lessons still to learn,” *BMJ Qual Saf*, vol. 27, no. 6, pp. 417–420, 2018, doi: 10.1136/bmjqs-2017-007563.

[115] N. M. Ivers *et al.*, “No more ‘business as usual’ with audit and feedback interventions: Towards an agenda for a reinvigorated intervention,” *Implementation Science*, vol. 9, no. 1, pp. 1–8, 2014, doi: 10.1186/1748-5908-9-14.

[116] KEMRI Wellcome Trust, “About CIN - The Clinical Information Network.” Accessed: Sep. 10, 2024. [Online]. Available: https://dodilion.com/cin/about-us/

[117] “Kemri Wellcome Trust Research Programme.” Accessed: Nov. 24, 2020. [Online]. Available: https://kemri-wellcome.org/

[118] M. English, P. Ayieko, R. Nyamai, F. Were, D. Githanga, and G. Irimu, “What do we think we are doing How might a clinical information network be promoting implementation of recommended paediatric care practices in Kenyan hospitals,” *Health Res Policy Syst*, vol. 15, no. 1, pp. 1–12, 2017, doi: 10.1186/s12961-017-0172-1.

[119] H. Pinnock *et al.*, “Standards for Reporting Implementation Studies (StaRI) Statement,” *BMJ (Online)*, vol. 356, no. March, pp. 1–9, 2017, doi: 10.1136/bmj.i6795.

[120] S. Theobald *et al.*, “Implementation research: new imperatives and opportunities in global health,” *The Lancet*, vol. 392, no. 10160, pp. 2214–2228, 2018, doi: 10.1016/S0140-6736(18)32205-0.

[121] B. Allemang, K. Sitter, and G. Dimitropoulos, “Pragmatism as a paradigm for patient-oriented research,” Feb. 01, 2022, *John Wiley and Sons Inc*. doi: 10.1111/hex.13384.

[122] M. Y. Feilzer, “Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm,” *J Mix Methods Res*, vol. 4, no. 1, pp. 6–16, Jan. 2010, doi: 10.1177/1558689809349691.

[123] A. Mitchell, “A Review of Mixed Methods, Pragmatism and Abduction Techniques,” *The Electronic Journal of Business Research Methods*, vol. 16, no. 3, pp. 103–116, 2018, [Online]. Available: www.ejbrm.com

[124] K. M. Long, F. McDermott, and G. N. Meadows, “Being pragmatic about healthcare complexity: Our experiences applying complexity theory and pragmatism to health services research,” *BMC Med*, vol. 16, no. 1, Jun. 2018, doi: 10.1186/s12916-018-1087-6.

[125] G. J. Burkholder, K. A. Cox, L. M. Crawford, and J. Hitchcock, *Research design and methods : an applied guide for the scholar-practitioner*. Los Angeles: SAGE Publications, Inc, 2020.

[126] C. Marshall, G. B. Rossman, and G. L. Blanco, *Designing qualitative research*, Seventh edition. Los Angeles: SAGE, 2022.

[127] S. M. Ravitch and M. Riggan, *Reason & rigor : how conceptual frameworks guide research*, Second edition. Thousand Oaks: SAGE Publications, Inc, 2016.

[128] J. W. Creswell and V. L. Plano Clark, *Designing and conducting mixed methods research*, Third edition. Los Angeles: SAGE, 2017.

[129] Government of Kenya - Ministry of Health, “Kenya Health Workforce Report: The Status of Healthcare Professionals in Kenya, 2015,” 2015. [Online]. Available: http://www.health.go.ke

[130] by Jalemba Aluvaala *et al.*, "Delivery outcomes and patterns of morbidity and mortality for neonatal admissions in five Kenyan hospitals", doi: 10.1093/tropej/fmv024.

[131] KEMRI Wellcome Trust, "CIN – The Clinical Information Network." Accessed: Nov. 03, 2023. [Online]. Available: https://cin.kemri-wellcome.org/

[132] "Kemri Wellcome Trust Research Programme." Accessed: Nov. 24, 2020. [Online]. Available: https://kemri-wellcome.org/

[133] M. English *et al.*, "Employing learning health system principles to advance research on severe neonatal and paediatric illness in Kenya," Mar. 23, 2021, *BMJ Publishing Group*. doi: 10.1136/bmjgh-2021-005300.

[134] T. L. Beauchamp and J. F. Childless, "Principles of Biomedical Ethics," *Int Clin Psychopharmacol*, vol. 6 (2), pp. 129–130, 1991.

[135] The EQUATOR Network, "Enhancing the QUAlity and Transparency Of health Research." Accessed: Sep. 09, 2024. [Online]. Available: https://www.equator-network.org

[136] A. C. Tricco *et al.*, "PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation," *Ann Intern Med*, vol. 169, no. 7, pp. 467–473, 2018, doi: 10.7326/M18-0850.

[137] A. Tong, P. Sainsbury, and J. Craig, "Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups," *International Journal for Quality in Health Care*, vol. 19, no. 6, pp. 349–357, 2007, doi: 10.1093/intqhc/mzm042.

[138] Critical Appraisal Skills Programme, "CASP (Critical Appraisal Skills Programme) checklist." Accessed: Sep. 09, 2024. [Online]. Available: https://casp-uk.net/casp-tools-checklists/

[139] "Oxford Tropical Research Ethics Committee (OxTREC) | Research Support." Accessed: Sep. 27, 2024. [Online]. Available: https://researchsupport.admin.ox.ac.uk/governance/ethics/committees/oxtrec

[140] D. P. Kaplan, Robert S.; Norton, “Using the Balanced Scorecard as a Strategic Management System,” *Harv Bus Rev*, vol. 88, no. 2, p. 534, 2007, doi: 10.1016/j.fertnstert.2007.05.031.

[141] Stephen. Few, “Dashboard Confusion Revisited,” *Perceptual Edge*, pp. 1–4, 2004.

[142] S. Few, *Information Dashboard Design - Displaying data for at-a-glance monitoring*, Second Edi. Burlingame, California, USA: Analytics Press, 2013.

[143] WHO-UNICEF Expert and Country Consultation on Small and/or Sick Newborn Care Group, “A comprehensive model for scaling up care for small and/or sick newborns at district level-based on country experiences presented at a WHO-UNICEF expert consultation,” 2023. Accessed: Sep. 05, 2024. [Online]. Available: https://jogh.org/wp-content/uploads/2023/04/jogh-13-03023.pdf

[144] B. C. Wallace, K. Small, C. E. Brodley, J. Lau, and T. A. Trikalinos, “Deploying an interactive machine learning system in an evidence-based practice center: Abstrackr.,” in *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium (IHI ’12),* , ACM Digital Library, 2013, pp. 819-824.

[145] J. Cohen, “A COEFFICIENT OF AGREEMENT FOR NOMINAL SCALES 1.”

[146] The World Bank, “World Bank Country and Lending Groups.” [Online]. Available: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519

[147] BORN Ontario, “The Better Outcomes Registry & Network (BORN) Ontario.” [Online]. Available: https://www.bornontario.ca/Modules/contact/search.aspx?s=gzyzez7QxwumqjwReGz9VgeQuAleQuAl

[148] C. Flannigan, U. Toland, and M. Hogan, “Clinical Audit Dovepress Dashboards in neonatology,” 2010. [Online]. Available: www.dovepress.com

[149] A. E. Sprague *et al.*, “Measuring Quality in Maternal-Newborn Care: Developing a Clinical Dashboard,” *Journal of Obstetrics and Gynaecology Canada*, vol. 35, no. 1, pp. 29–38, 2013, doi: 10.1016/S1701-2163(15)31045-8.

[150] R. A. Simms, H. Ping, A. Yelland, A. J. Beringer, R. Fox, and T. J. Draycott, “Development of maternity dashboards across a UK health region; Current practice, continuing problems,” *European Journal of Obstetrics and Gynecology and Reproductive Biology*, vol. 170, no. 1, pp. 119–124, 2013, doi: 10.1016/j.ejogrb.2013.06.003.

[151] J. Gaspar, J. Chagas, G. C. Osanan, R. Cruz-Correa, and Z. S. N. Reis, “Maternal and neonatal healthcare information system: Development of an obstetric electronic health record and healthcare indicators dashboard,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8060 LNCS, pp. 62–76, 2013, doi: 10.1007/978-3-642-40093-3_5.

[152] S. Bhardwaj *et al.*, “Elimination of mother-to-child transmission of HIV in South Africa: Rapid scale-up using quality improvement,” *South African Medical Journal*, vol. 104, no. 3, pp. 239–243, 2014, doi: 10.7196/SAMJ.7605.

[153] M. S. Raban, C. Bamford, Y. Joolay, and M. C. Harrison, “Impact of an educational intervention and clinical performance dashboard on neonatal bloodstream infections,” *South African Medical Journal*, vol. 105, no. 7, pp. 564–566, 2015, doi: 10.7196/SAMJnew.7764.

[154] A. Temko, W. Marnane, G. Boylan, and G. Lightbody, “Clinical implementation of a neonatal seizure detection algorithm,” *Decis Support Syst*, vol. 70, pp. 86–96, 2015, doi: 10.1016/j.dss.2014.12.006.

[155] S. Dunn *et al.*, “A mixed methods evaluation of the maternal-newborn dashboard in Ontario: Dashboard attributes, contextual factors, and facilitators and barriers to use: A study protocol,” *Implementation Science*, vol. 11, no. 1, pp. 1–11, 2016, doi: 10.1186/s13012-016-0427-1.

[156] D. Weiss *et al.*, “Effect of a population-level performance dashboard intervention on maternal-newborn outcomes: An interrupted time series study,” *BMJ Qual Saf*, vol. 27, no. 6, pp. 425–436, 2018, doi: 10.1136/bmjqs-2017-007361.

[157] D. Díaz, “Visual tools for the exploration of growth data in a cohort of kangaroo,” no. 1, pp. 9–16, 2017.

[158] L. Szymanski, C. Arnold, A. J. Vaught, S. LaMantia, T. Harris, and A. J. Satin, “Implementation of a multicenter shoulder dystocia injury prevention program,” *Semin Perinatol*, vol. 41, no. 3, pp. 187–194, 2017, doi: 10.1053/j.semperi.2017.03.008.

[159] C. Patrick, H. Leonard, W. Novicoff, R. West, C. Crews, and W. Andrews, “Improving analysis and reporting for Virginia’s Division of Consolidated Laboratory Services newborn screening program,” *2017 Systems and Information Engineering Design Symposium, SIEDS 2017*, pp. 237–242, 2017, doi: 10.1109/SIEDS.2017.7937723.

[160] J. Reszel *et al.*, “Use of a maternal newborn audit and feedback system in Ontario: A collective case study,” *BMJ Qual Saf*, vol. 28, no. 8, pp. 635–644, 2019, doi: 10.1136/bmjqs-2018-008354.

[161] P. Rezaei-hachesu *et al.*, “The design and evaluation of an antimicrobial resistance surveillance system for neonatal intensive care units in Iran,” *Int J Med Inform*, vol. 115, no. April, pp. 24–34, 2018, doi: 10.1016/j.ijmedinf.2018.04.007.

[162] K. J. Bunch, B. Allin, M. Jolly, T. Hardie, and M. Knight, “Developing a set of consensus indicators to support maternity service quality improvement: using Core Outcome Set methodology including a Delphi process,” *BJOG*, vol. 125, no. 12, pp. 1612–1618, 2018, doi: 10.1111/1471-0528.15282.

[163] M. M. Jankowska *et al.*, “An online geographic data visualization tool to relate preterm births to environmental factors,” *Prev Chronic Dis*, vol. 16, no. 8, pp. 1–9, 2019, doi: 10.5888/pcd16.180498.

[164] M. S. Patel, B. Rathi, K. Tashfeen, and M. A. Yarubi, “Development and implementation of maternity dashboard in regional hospital for quality improvement at ground level: A pilot study,” *Oman Med J*, vol. 34, no. 3, pp. 194–199, 2019, doi: 10.5001/omj.2019.38.

[165] T. Begum *et al.*, “Perceptions and experiences with district health information system software to collect and utilize health data in Bangladesh: A qualitative exploratory

study,” *BMC Health Serv Res*, vol. 20, no. 1, pp. 1–14, 2020, doi: 10.1186/s12913-020-05322-2.

[166] S. Arulkumaran, E. Chandraharan, T. Mahmood, O. Louca, and C. Mannion, “Maternity dashboard: clinical performance and governance score card,” *Good Practice*, no. 7, pp. 1–8, 2008.

[167] A. E. Sprague *et al.*, “Measuring Quality in Maternal-Newborn Care: Developing a Clinical Dashboard,” *Journal of Obstetrics and Gynaecology Canada*, vol. 35, no. 1, pp. 29–38, 2013, doi: 10.1016/S1701-2163(15)31045-8.

[168] K. Schwaber and J. Sutherland, “The Scrum Guide The Definitive Guide to Scrum: The Rules of the Game,” 2020.

[169] M. Hicks and J. S. Foster, “Adapting Scrum to Managing a Research Group *,” 2010. [Online]. Available: http://www.controlchaos.com/images/diagram/flow.gif

[170] T. Munzner and E. Maguire, *Visualization analysis & design*, 1st edition. in A K Peters visualization series. Boca Raton, Florida: CRC Press, 2015.

[171] J. Bernard *et al.*, “VisInfo: a digital library system for time series research data based on exploratory search—a user-centered design approach,” *International Journal on Digital Libraries*, vol. 16, no. 1, pp. 37–59, Apr. 2015, doi: 10.1007/s00799-014-0134-y.

[172] M. McGuire, A. Gangopadhyay, A. Komlodi, and C. Swan, “A user-centered design for a spatial data warehouse for data exploration in environmental research,” *Ecol Inform*, vol. 3, no. 4–5, pp. 273–285, Oct. 2008, doi: 10.1016/j.ecoinf.2008.08.002.

[173] R. E. Roth, K. S. Ross, and A. M. MacEachren, “User-centered design for interactive maps: A case study in crime analysis,” *ISPRS Int J Geoinf*, vol. 4, no. 1, pp. 262–301, Mar. 2015, doi: 10.3390/ijgi4010262.

[174] V. Janakiraman and J. Ecker, “Quality in Obstetric Care: Measuring What Matters,” *Obstetrics & Gynecology*, vol. 116, no. 3, pp. 728–732, Sep. 2010, doi: 10.1097/AOG.0b013e3181ea4d4f.

[175] R. C. Hermann and R. H. Palmer, "Common Ground: A Framework for Selecting Core Quality Measures for Mental Health and Substance Abuse Care," *Psychiatric Services*, vol. 53, no. 3, pp. 281–287, Mar. 2002, doi: 10.1176/appi.ps.53.3.281.

[176] D. Pencheon, "The good indicator guide: understanding how to use and choose indicators," *NHS Institute for Innovation and Improvement*, 2008.

[177] I. El Azami, M. O. Cherkaoui Malki, and C. Tahon, "Integrating hospital information systems in healthcare institutions: A mediation architecture," *J Med Syst*, vol. 36, no. 5, pp. 3123–3134, Oct. 2012, doi: 10.1007/s10916-011-9797-8.

[178] "National Neonatal Audit Programme - Data dashboard | RCPCH." Accessed: Sep. 04, 2024. [Online]. Available: https://www.rcpch.ac.uk/resources/nnap-data-dashboard

[179] M. Cumpston and E. Flemyng, *Chapter IV: Updating a review [last updates August 2023]. In: Cochrane Handbook for Systematic Reviews of Interventions*, vol. 6.5. Wiley, 2019. doi: 10.1002/9781119536604.

[180] S. Few, "Data Visualization Effectiveness Profile," *Perceptual Edge*, no. March, pp. 1–11, 2017, [Online]. Available: https://www.perceptualedge.com/articles/visual_business_intelligence/data_visualization_effectiveness_profile.pdf

[181] A. Rind *et al.*, "Interactive information visualization to explore and query electronic health records," *Foundations and Trends in Human-Computer Interaction*, vol. 5, no. 3, pp. 207–298, 2011, doi: 10.1561/1100000039.

[182] V. Saltenis and J. Aušraite, "Data visualization: ideas, methods, and problems," *Informatics in education*, vol. 1, pp. 129–148, 2002, doi: 10.15388/INFEDU.2002.11.

[183] S. Franconeri, L. M. K. Padilla, P. Shah, J. M. Zacks, and J. Hullman, "The Science of Visual Data Communication: What Works," *Psychological Science in the Public Interest*, vol. 22, pp. 110–161, 2021, doi: 10.1177/15291006211051956.

[184] J. Wagemans *et al.*, "A century of Gestalt psychology in visual perception: II. Conceptual and theoretical foundations," *Psychol Bull*, vol. 138, no. 6, pp. 1218–1252, Nov. 2012, doi: 10.1037/a0029334.

[185] T. Zuk, L. Schlesier, P. Neumann, M. S. Hancock, and S. Carpendale, "Heuristics for information visualization evaluation," *Proceedings of BELIV'06: BEyond time and errors - novel EvaLuation methods for Information Visualization. A workshop of the AVI 2006 International Working Conference*, pp. 1–6, 2006, doi: 10.1145/1168149.1168162.

[186] E. Hehman and S. Y. Xie, "Doing Better Data Visualization," *Adv Methods Pract Psychol Sci*, vol. 4, p., 2021, doi: 10.1177/25152459211045334.

[187] Y. Zhu, "LNCS 4842 - Measuring Effective Data Visualization," 2007.

[188] ISO Online Browsing Platform (OBP), "ISO 9241-11:2018(en) Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts," https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en.

[189] "User Satisfaction vs. Performance Metrics." Accessed: Nov. 05, 2023. [Online]. Available: https://www.nngroup.com/articles/satisfaction-vs-performance-metrics/

[190] "First Rule of Usability? Don't Listen to Users." Accessed: Dec. 20, 2023. [Online]. Available: https://www.nngroup.com/articles/first-rule-of-usability-dont-listen-to-users/

[191] R. L. Sinkowitz-Cochran, "Survey Design: To Ask or Not to Ask? That is the Question...," *Clinical Infectious Diseases*, vol. 56, no. 8, pp. 1159–64, 2013, doi: 10.1093/cid/cit005.

[192] "Should You Use 5 Or 7 Point Scales? – MeasuringU." Accessed: Dec. 20, 2023. [Online]. Available: https://measuringu.com/scale-points/

[193] J. C. F. De Winter and D. Dodou, "Five-Point Likert Items: t test versus Mann-Whitney-Wilcoxon," vol. 15, no. 11, 2010.

[194] G. Norman, "Likert scales, levels of measurement and the 'laws' of statistics," *Advances in Health Sciences Education*, vol. 15, no. 5, pp. 625–632, Dec. 2010, doi: 10.1007/s10459-010-9222-y.

[195] B. P. Subedi, "Using Likert Type Data in Social Science Research: Confusion, Issues and Challenges," 2016. [Online]. Available: www.ijcas.net

[196] G. M. Sullivan and A. R. Artino, "Analyzing and Interpreting Data From Likert-Type Scales," *J Grad Med Educ*, vol. 5, no. 4, pp. 541–542, Dec. 2013, doi: 10.4300/jgme-5-4-18.

[197] A. Cazañas, A. de San Miguel, and E. Parra, “Estimating Sample Size for Usability Testing (Estimación del tamaño de la muestra para pruebas de usabilidad),” *Enfoque UTE*, no. 1, pp. 172–185, 2017, [Online]. Available: http://ingenieria.ute.edu.ec/enfoqueute/

[198] C. Barnum, N. Bevan, G. Cockton, and J. N. Nielsen, “The ‘Magic Number 5’: Is It Enough for Web Testing?” [Online]. Available: www.useit.com/alertbox/2000319.html

[199] Budiu Raluca and Moran Kate, “How Many Participants for Quantitative Usability Studies: A Summary of Sample-Size Recommendations.” Accessed: Jan. 14, 2024. [Online]. Available: https://www.nngroup.com/articles/summary-quant-sample-sizes/

[200] L. Faulkner, “Beyond the five-user assumption: Benefits of increased sample sizes in usability testing,” 2003.

[201] J. Sauro, “How to Find the Sample Size for 8 Common Research Designs – MeasuringU.” Accessed: Jan. 14, 2024. [Online]. Available: https://measuringu.com/sample-size-designs/

[202] R Core Team, “R: A language and environment for statistical computing.,” 2020, *R Foundation for Statistical Computing, Vienna, Austria*.

[203] IBM Corp., “IBM SPSS Statistics for Macintosh, Version 29.0.” Accessed: Oct. 07, 2024. [Online]. Available: https://www.ibm.com/support/pages/how-cite-ibm-spss-statistics-or-earlier-versions-spss

[204] D. Muijs, “Doing Quantitative Research in Education with SPSS - Daniel Muijs - Google Books,” 2012.

[205] N. Schwarz, “Self-Reports_How the Questions Shape the Answers,” 2019.

[206] M. Soegaard, “What is Eye Tracking in UX? - Interaction Design Foundation.” Accessed: Aug. 02, 2024. [Online]. Available: https://www.interaction-design.org/literature/article/eye-tracking-ux

[207] T. I. Wanvik, “Reflections on the Research Process,” in *Contested Energy Spaces: Disassembling Energyscapes of the Canadian North*, Cham: Springer International Publishing, 2019, pp. 17–42. doi: 10.1007/978-3-030-02396-6_3.

[208] G. Payne and M. Williams, “Generalization in qualitative research,” *Sociology*, vol. 39, no. 2, pp. 295–314, 2005, doi: 10.1177/0038038505050540.

[209] VERBI Software, “MAXQDA 2022 [computer software],” 2021, *Available from maxqda.com, Berlin, Germany*.

[210] L. Prior, *The Oxford Handbook of Qualitative Research - Chapter 19 Content Analysis*. Oxford University Press, 2020. doi: 10.1093/oxfordhb/9780190847388.001.0001.

[211] N. K. Gale, G. Heath, E. Cameron, S. Rashid, and S. Redwood, “Using the framework method for the analysis of qualitative data in multi-disciplinary health research,” *BMC Med Res Methodol*, vol. 13, no. 1, 2013, doi: 10.1186/1471-2288-13-117.

[212] N. Bergen and R. Labonté, “‘Everything Is Perfect, and We Have No Problems’: Detecting and Limiting Social Desirability Bias in Qualitative Research,” *Qual Health Res*, vol. 30, no. 5, pp. 783–792, Apr. 2020, doi: 10.1177/1049732319889354.

[213] L. A. Palinkas, S. M. Horwitz, C. A. Green, J. P. Wisdom, N. Duan, and K. Hoagwood, “Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research,” *Administration and Policy in Mental Health and Mental Health Services Research*, vol. 42, no. 5, pp. 533–544, Sep. 2015, doi: 10.1007/s10488-013-0528-y.

[214] S. Campbell *et al.*, “Purposive sampling: complex or simple? Research case examples,” *Journal of Research in Nursing*, vol. 25, no. 8, pp. 652–661, Dec. 2020, doi: 10.1177/1744987120927206.

[215] J. Noyes, A. Booth, G. Moore, K. Flemming, Ö. Tunçalp, and E. Shakibazadeh, “Synthesising quantitative and qualitative evidence to inform guidelines on complex interventions: Clarifying the purposes, designs and outlining some methods,” *BMJ Glob Health*, vol. 4, no. Supplement1, 2019, doi: 10.1136/bmjgh-2018-000893.

[216] M. J. Johnson and C. R. May, “Promoting professional behaviour change in healthcare: what interventions work, and why? A theory-led overview of systematic reviews,” *Open*, vol. 5, p. 8592, 2015, doi: 10.1136/bmjopen-2015.

[217] A. A. Powell *et al.*, “More than a score: a qualitative study of ancillary benefits of performance measurement”, doi: 10.1136/bmjqs.

[218] M. Gort, M. Broekhuis, and G. Regts, “How teams use indicators for quality improvement - A multiple-case study on the use of multiple indicators in multidisciplinary breast cancer teams,” *Soc Sci Med*, vol. 96, pp. 69–77, Nov. 2013, doi: 10.1016/j.socscimed.2013.06.001.

[219] M. Cameron, G. Penney, G. MacLennan, S. McLeer, and A. Walker, “Impact on maternity professionals of novel approaches to clinical audit feedback,” *Eval Health Prof*, vol. 30, no. 1, pp. 75–95, Mar. 2007, doi: 10.1177/0163278706297337.

[220] C. Palmer, J. Bycroft, K. Healey, A. Field, and M. Ghafel, “Can formal collaborative methodologies improve quality in primary health care in New Zealand? insights from the EQUIPPED Auckland Collaborative,” *J Prim Health Care*, vol. 4, no. 4, pp. 328–336, 2012, doi: 10.1071/hc12328.

[221] V. L. Payne and S. J. Hysong, “Model depicting aspects of audit and feedback that impact physicians’ acceptance of clinical performance feedback,” *BMC Health Serv Res*, vol. 16, no. 1, Jul. 2016, doi: 10.1186/s12913-016-1486-3.

[222] L. M. Chadwick, A. Macphail, J. E. Ibrahim, L. Mcauliffe, S. Koch, and Y. Wells, “Senior staff perspectives of a quality indicator program in public sector residential aged care services: A qualitative cross-sectional study in Victoria, Australia,” *Australian Health Review*, vol. 40, no. 1, pp. 54–62, 2016, doi: 10.1071/AH14197.

[223] R. Y. Wang and D. M. Strong, “Beyond Accuracy: What Data Quality Means to Data Consumers,” 1996.

[224] K.-U. Sattler, “Data Quality Dimensions,” *Encyclopedia of Database Systems*, pp. 1–5, 2016, doi: 10.1007/978-1-4899-7993-3_108-2.

[225] L. Jeffs, S. Beswick, J. Lo, Y. Lai, A. Chhun, and H. Campbell, “Insights from staff nurses and managers on unit-specific nursing performance dashboards: A qualitative study,” *BMJ Qual Saf*, vol. 23, no. 12, pp. 1001–1006, Dec. 2014, doi: 10.1136/bmjqs-2013-002595.

[226] R. Mannion and M. Goddard, “Impact of published clinical outcomes data: case study in NHS hospital trusts,” 1993.

[227] S. G. Yi *et al.*, “Surgeon-specific performance reports in general surgery: An observational study of initial implementation and adoption,” *J Am Coll Surg*, vol. 217, no. 4, 2013, doi: 10.1016/j.jamcollsurg.2013.04.040.

[228] Y. Tanahashi, N. Leaf, and K. Ma, “A Study On Designing Effective Introductory Materials for Information Visualization,” *Computer Graphics Forum*, vol. 35, p., 2016, doi: 10.1111/cgf.13009.

[229] T. Greenhalgh, G. Robert, P. Bate, O. Kyriakidou, and F. Macfarlane, “How to Spread Good Ideas A systematic review of the literature on diffusion, dissemination and sustainability of innovations in health service delivery and organisation Report for the National Co-ordinating Centre for NHS Service Delivery and Organisation R & D (NCCSDO),” 2004.

[230] N. Muinga *et al.*, “Digital health Systems in Kenyan Public Hospitals: A mixed-methods survey,” Jan. 06, 2020, *BioMed Central Ltd*. doi: 10.1186/s12911-019-1005-7.

[231] C. J. Lauw, J. Rahman, A. Brankovic, M. Tracy, and S. Khanna, “Development of an Interactive Dashboard to Analyse Physiological Signals in the Neonatal Intensive Care Unit,” in *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, IEEE, Jul. 2023, pp. 1–4. doi: 10.1109/EMBC40787.2023.10340576.

[232] BORN Ontario, “The Better Outcomes Registry & Network (BORN) Ontario.” [Online]. Available: https://www.bornontario.ca/Modules/contact/search.aspx?s=gzyzez7QxwumqjwReGz9VgeQuAleQuAl

[233] S. C. Buttigieg, A. Pace, and C. Rathert, “Hospital performance dashboards: a literature review,” *J Health Organ Manag*, vol. 31, no. 3, pp. 385–406, 2017, doi: 10.1108/JHOM-04-2017-0088.

[234] M. English, P. Ayieko, R. Nyamai, F. Were, D. Githanga, and G. Irimu, "What do we think we are doing How might a clinical information network be promoting implementation of recommended paediatric care practices in Kenyan hospitals," *Health Res Policy Syst*, vol. 15, no. 1, pp. 1–12, 2017, doi: 10.1186/s12961-017-0172-1.

[235] "The Extreme Presentation Method." Accessed: Sep. 22, 2024. [Online]. Available: https://extremepresentation.com/

[236] J. Nielsen and Nielsen Norman Groups, "Iterative Design of User Interfaces." Accessed: Oct. 03, 2024. [Online]. Available: https://www.nngroup.com/articles/iterative-design/

[237] S. Tselova, "Introducing scrollytelling: the future of immersive data storytelling," May 10, 2023.

[238] C. Stoiber *et al.*, "Comparative Evaluations of Visualization Onboarding Methods," *Vis. Informatics*, vol. 6, pp. 34–50, 2022, doi: 10.48550/arXiv.2203.15418.

[239] A. Srinivasan, S. Drucker, A. Endert, and J. Stasko, "Augmenting Visualizations with Interactive Data Facts to Facilitate Interpretation and Communication," *IEEE Trans Vis Comput Graph*, vol. 25, pp. 672–681, 2019, doi: 10.1109/TVCG.2018.2865145.

[240] S. Fleury and N. Chaniaud, "Multi-user centered design: acceptance, user experience, user research and user testing," *Theor Issues Ergon Sci*, vol. 25, no. 2, pp. 209–224, 2024, doi: 10.1080/1463922X.2023.2166623.

[241] K. Whitenton and Nielsen Norman Group, "Unmoderated User Tests: How and Why to Do Them." Accessed: Oct. 08, 2024. [Online]. Available: https://www.nngroup.com/articles/unmoderated-usability-testing/

[242] K. L. Schwartz *et al.*, "Best practice guidance for antibiotic audit and feedback interventions in primary care: a modified Delphi study from the Joint Programming Initiative on Antimicrobial resistance: Primary Care Antibiotic Audit and Feedback Network (JPIAMR-PAAN)," *Antimicrob Resist Infect Control*, vol. 12, no. 1, p. 72, Jul. 2023, doi: 10.1186/s13756-023-01279-z.

[243] M. Lazzerini, C. Semenzato, J. Kaur, B. Covi, and G. Argentini, "Women's suggestions on how to improve the quality of maternal and newborn hospital care: A qualitative study

in Italy using the WHO standards as framework for the analysis," *BMC Pregnancy Childbirth*, vol. 20, no. 1, Apr. 2020, doi: 10.1186/s12884-020-02893-0.

[244] C. Horwood, L. Haskins, S. Luthuli, and N. McKerrow, "Communication between mothers and health workers is important for quality of newborn care: A qualitative study in neonatal units in district hospitals in South Africa," *BMC Pediatr*, vol. 19, no. 1, Dec. 2019, doi: 10.1186/s12887-019-1874-z.

[245] C. Gribbin *et al.*, "Exploring the influence of postnatal depression on neonatal care practices among mothers in Western Kenya: A qualitative study," *Women's Health*, vol. 19, Jan. 2023, doi: 10.1177/17455057231189547.

[246] L. Moore *et al.*, "Economic evaluations of audit and feedback interventions: a systematic review," Jun. 24, 2022, *BMJ Publishing Group*. doi: 10.1136/bmjqs-2022-014727.

[247] A. I. Resource and J. Perlo, "IHI Framework for Improving Joy in Work," 2017.

[248] A. Lawson-Body, L. Lawson-Body, and A. Illia, "Data Visualization: Developing and Validating Dashboard Measurement Instruments," *Journal of Computer Information Systems*, vol. 63, no. 3, pp. 479–491, 2023, doi: 10.1080/08874417.2022.2073295.

[249] R. Sawatzky *et al.*, "Implications of response shift for micro-, meso-, and macro-level healthcare decision-making using results of patient-reported outcome measures," *Quality of Life Research*, vol. 30, no. 12, pp. 3343–3357, Dec. 2021, doi: 10.1007/s11136-021-02766-9.

[250] H. C. Lim *et al.*, "Toward a Learning Health Care System: A Systematic Review and Evidence-Based Conceptual Framework for Implementation of Clinical Analytics in a Digital Hospital," Mar. 01, 2022, *Georg Thieme Verlag*. doi: 10.1055/s-0042-1743243.

[251] N. S. Riegels, ; Lindsay, and A. Mazotti, "Quality and Health Equity Dashboards for Internal Medicine Residents: Interactive Displays to Promote Systems-Based Practice and Practice-Based Learning and Improvement," 2023.

[252] The Health Foundation, "Untapped potential: Investing in health and care data analytics." Accessed: Oct. 03, 2024. [Online]. Available: https://reader.health.org.uk/untapped-potential/summary

[253] World Health Organization, “SURVIVE AND THRIVE: Transforming care for every small and sick newborn,” Chapter 5, Figure 5.1, 2019. [Online]. Available: https://apps.who.int/iris/bitstream/handle/10665/326495/9789241515887-eng.pdf?ua=1

[254] L. Hui *et al.*, “Collaborative maternity and newborn dashboard (CoMaND) for the COVID-19 pandemic: A protocol for timely, adaptive monitoring of perinatal outcomes in Melbourne, Australia,” *BMJ Open*, vol. 11, no. 11, Nov. 2021, doi: 10.1136/bmjopen-2021-055902.

[255] M. S. Q. Murphy *et al.*, “Data Resource Profile: Better Outcomes Registry & Network (BORN) Ontario,” *Int J Epidemiol*, vol. 50, no. 5, pp. 1416-1417H, Oct. 2021, doi: 10.1093/ije/dyab033.

[256] E. K. Wilkinson *et al.*, “Reactions to the use of evidence-based performance indicators in primary care: a qualitative study.” [Online]. Available: www.qualityhealthcare.com

[257] C. Nessim, C. M. Bensimon, B. Hales, C. Laflamme, D. Fenech, and A. Smith, “Surgical site infection prevention: A qualitative analysis of an individualized audit and feedback model,” *J Am Coll Surg*, vol. 215, no. 6, pp. 850–857, Dec. 2012, doi: 10.1016/j.jamcollsurg.2012.08.007.

[258] T. L. Guldberg, P. Vedsted, T. Lauritzen, and V. Zoffmann, “Suboptimal quality of type 2 diabetes care discovered through electronic feedback led to increased nurse–GP cooperation. A qualitative study,” *Prim Care Diabetes*, vol. 4, no. 1, pp. 33–39, Apr. 2010, doi: 10.1016/j.pcd.2010.01.001.

[259] M. Exworthy *et al.*, “The role of performance indicators in changing the autonomy of the general practice profession in the UK,” 2003.

[260] K. M. Cresswell *et al.*, “An embedded longitudinal multi-faceted qualitative evaluation of a complex cluster randomized controlled trial aiming to reduce clinically important errors in medicines management in general practice,” *Trials*, vol. 13, Jun. 2012, doi: 10.1186/1745-6215-13-78.

[261] S. J. Hysong, M. K. Knox, and P. Haidet, “Examining clinical performance feedback in patient-aligned care teams,” *J Gen Intern Med*, vol. 29, no. SUPPL. 2, 2014, doi: 10.1007/s11606-013-2707-7.

[262] C. Morrell, G. Harvey, and A. Kitson, “Practitioner based quality improvement: a review of the Royal College of Nursing’s dynamic standard setting system,” 1997.

[263] A. M. Grant, B. Guthrie, and T. Dreischulte, “Developing a complex intervention to improve prescribing safety in primary care: mixed methods feasibility and optimisation pilot study,” *BMJ Open*, vol. 4, p. 4153, 2014, doi: 10.1136/bmjopen-2013.

[264] N. Ivers *et al.*, “Perceived discordance between population-level quality targets and patient-centred care.” [Online]. Available: www.cfp.ca.

[265] M. Dixon-Woods, S. Redwood, M. Leslie, J. Minion, G. P. Martin, and J. J. Coleman, “Improving quality and safety of care using technovigilance: An ethnographic case study of secondary use of data from an electronic prescribing and decision support system,” *Milbank Quarterly*, vol. 91, no. 3, pp. 424–454, Sep. 2013, doi: 10.1111/1468-0009.12021.

[266] E. Nouwens, J. Van Lieshout, and M. Wensing, “Determinants of impact of a practice accreditation program in primary care: A qualitative study,” *BMC Fam Pract*, vol. 16, no. 1, Jul. 2015, doi: 10.1186/s12875-015-0294-x.

[267] A. A. Powell *et al.*, “Unintended consequences of implementing a national performance measurement system into local practice,” *J Gen Intern Med*, vol. 27, no. 4, pp. 405–412, Apr. 2012, doi: 10.1007/s11606-011-1906-3.

[268] B. Vachon *et al.*, “Engaging primary care practitioners in quality improvement: Making explicit the program theory of an interprofessional education intervention,” *BMC Health Serv Res*, vol. 13, no. 1, 2013, doi: 10.1186/1472-6963-13-106.

[269] S. Tierney, R. Kislov, and C. Deaton, “A qualitative study of a primary-care based intervention to improve the management of patients with heart failure: The dynamic relationship between facilitation and context,” *BMC Fam Pract*, vol. 15, no. 1, Sep. 2014, doi: 10.1186/1471-2296-15-153.

[270] O. Groene, N. Klazinga, V. Kazandjian, P. Lombrail, and P. Bartels, “The World Health Organization Performance Assessment Tool for quality improvement in hospitals (PATH): An analysis of the pilot implementation in 37 hospitals,” *International Journal for Quality in Health Care*, vol. 20, no. 3, pp. 155–161, Jun. 2008, doi: 10.1093/intqhc/mzn010.

[271] S. Turner, J. Higginson, C. A. Oborne, R. E. Thomas, A. I. G. Ramsay, and N. J. Fulop, “Codifying knowledge to improve patient safety: A qualitative study of practice-based interventions,” *Soc Sci Med*, vol. 113, pp. 169–176, 2014, doi: 10.1016/j.socscimed.2014.05.031.

[272] R. Mcdonald and M. Roland, “Pay for performance in primary care in England and california: Comparison of unintended consequences,” *Ann Fam Med*, vol. 7, no. 2, pp. 121–127, 2009, doi: 10.1370/afm.946.

[273] “HTML vs XML - javatpoint.” Accessed: Oct. 09, 2024. [Online]. Available: https://www.javatpoint.com/html-vs-xml