BMJ Open

Investigating the characteristics of health-related data collection tools used in randomised controlled trials in low-income and middle-income countries: protocol for a systematic review

Rithvik Karthikeyan ,[1] Noor Al-Shamaa ,[1] Eve Jacqueline Kelly,[1] Patrick Henn,[2] Frances Shiely,[3] Titus Divala ,[4] Kayode Philip Fadahunsi ,[5] John O'Donoghue [6,7]

To cite: Karthikeyan R, Al-Shamaa N, Kelly EJ, *et al*. Investigating the characteristics of health-related data collection tools used in randomised controlled trials in low-income and middle-income countries: protocol for a systematic review. *BMJ Open* 2024;**14**:e077148. doi:10.1136/bmjopen-2023-077148

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2023-077148).

Received 26 June 2023
Accepted 09 January 2024

For numbered affiliations see end of article.

Correspondence to
Dr Kayode Philip Fadahunsi; k.fadahunsi14@imperial.ac.uk

ABSTRACT

Introduction Health-related data collection tools, including digital ones, have become more prevalent across clinical studies in the last number of years. However, using digital data collection tools in low-income and middle-income countries presents unique challenges. In this review, we aim to provide an overview of the data collection tools currently being used in randomised controlled trials (RCTs) conducted in low-resource settings and evaluate the tools based on the characteristics outlined in the modified Mobile Survey Tool framework. These include functionality, reliability, usability, efficiency, maintainability, portability, effectiveness, cost–benefit, satisfaction, freedom from risk and context coverage. This evidence may provide a guide to selecting a suitable data collection tool for researchers planning to conduct research in low-income and middle-income countries for future studies.

Methods and analysis Searches will be conducted in four electronic databases: PubMed, CINAHL, Web of Science and EMBASE. For inclusion, studies must be a RCT, mention a health-related data collection tool and conducted in a low- and middle-income country. Only studies with available full-text and written in English will be included. The search was restricted to studies published between January 2005 and June 2023. This systematic review will use the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) tool. Two review authors will screen the titles and abstracts of search results independently for inclusion. In the initial screening process, the full-text articles will be retrieved if the abstract contains limited information about the study. Disagreements will be resolved through discussion. If the disagreement cannot be resolved, a third author (JO'D) will adjudicate. The study selection process will be outlined in a PRISMA flow-diagram. Data will be analysed using a narrative synthesis approach. The included studies and their outcomes will be presented in a table.

Ethics and dissemination Formal ethical approval is not required as primary data will not be collected in this study. The findings from this systematic review will be published in a peer-reviewed journal.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ The review will use four well-established databases, which cover the primary area of interest.
⇒ The formulation and development of comprehensive search terms that cover the various types of data collection is challenging.
⇒ Studies that are not published in English will be excluded.

PROSPERO registration number CRD42023405738.

INTRODUCTION

Data collection tools are a key part of service delivery and medical research, as it is the means through which statistics on a microlevel and macrolevel are gathered in relation to the healthcare of the patient and or at the population level via public health. Data collection tools are defined as any instrument used by researchers and healthcare professionals to collect data ranging from paper questionnaires to peak flow metres.[1] The field of clinical medicine is unique in that countless specialised data collection tools exist. For instance, one specialised data collection tool is the blood glucometer, which was invented in the 1970s.[2] While this was initially only available to practitioners, over time, it became more portable and widely accessible to the general public. The data collection of blood glucose levels on a public scale has played a significant role in discovering new diabetes medications and in calibrating the management of diabetes.[2] Similar patterns exist with other types of health-related data collection tools, as they have been instrumental in positively impacting public health.

Over the last number of years, the development of new digital technologies has enabled researchers to collect data in a more effective and efficient manner.[3] Digital data collection tools such as Mobile Survey Tools (MST), apps, wearable devices, Artificial Intelligence (AI), video and audio analytical tools, and internet-of-things-based products are becoming more prevalent in clinical research. One particular benefit of digital data collection tools is that they enable researchers to undertake these processes digitally and remotely, without requiring the physical presence of the patient (eg, remote monitoring). Other benefits include their cost-effectiveness and time efficiency,[4] which is of particular importance in low-resource settings. More recently, the application of wearable devices was highlighted during the COVID-19 pandemic.[5] These tools collected data on a range of parameters such as 'pulse, physical activity and sleep' in order to calculate the regional probability of a COVID-19 outbreak.[5] Hence, this review will also focus on the emerging digital aspect of health-related data collection tools used in low- and middle-income countries (LMICs).

As the years progress, data collection in healthcare will keep increasing. An editorial from 2023 by Rahman claims machine learning and deep learning techniques will greatly increase the volume of medical data collected in the future. The editorial cites the PATINA decision support tool as an example of an intelligent monitoring system that can prevent the hospitalisation rates of frail older adults.[6] Social media can also be used to harness data on patients to improve their management. This can occur on an individual and population level.[7]

According to the WHO, developing research capacity in LMICs is one of the key ways to promote global health equality.[8] More specifically, it is recommended that governments of LMICs enact policies that incentivise health research, offer financial support for higher education research departments and promote research partnerships between research bodies, academia and health providers.[8] One article highlighted that identifying and improving pre-existing data collection tools in LMICs can be instrumental in saving lives, particularly in emergency departments.[9] LMICs, which include LMICs, are defined as those with a Gross National Income (GNI) per capita of below US$4255.[10]

The country classification by GNI per capita for 2023 is presented in table 1.

Table 1 Country classification by Gross National Income (GNI) per capita for the year 2023 according to the World Bank Atlas method[10]

Country classification	GNI per capita
Low income	US$1085 or less
Lower middle income	US$1086 and US$4255
Upper middle income	US$4256 and US$13205
High income	US$13205 or more

A randomised controlled trial (RCT) is a research methodology in which participants are randomly assigned to one of two or more clinical interventions.[11] RCTs are considered the most scientifically rigorous method of hypothesis testing available and are regarded as the gold-standard study design for evaluating the effectiveness of interventions.[11] The use of RCTs as a study design is becoming more prevalent in LMICs.[12] Therefore, conducting research focused on RCTs could offer valuable guidance for researchers using RCTs in similar settings in the future.

In this review, the MST framework will be used to assess the characteristics of the data collection tools.[13] While there are various evaluation frameworks for certain subtypes of digital data collection tools such as wearables[14] and apps,[15] an overarching framework for all data collection tools does not exist. It may be challenging to evaluate the wide variety of data collection tools available under one framework. While the MST framework is designed for evaluating MSTs, the characteristics within this framework provide a comprehensive assessment that may be applicable to other data collection tools. It contains the key characteristics that practitioners may consider when choosing a data collection tool for their research. Fisher *et al* define the function of MSTs, stating, 'MSTs allow users to gather and transmit field data in real-time, standardise data storage and management, automate routine analyses and visualise data'.[13] This is broadly the function of all data collection tools. Therefore, a framework used to evaluate the functions of an MST can be applicable when evaluating the function of all data collection tools. Based on other frameworks reviewed,[16 17] the MST has been modified to include the following additional criteria:

- Type of data collection tool.
- Online or offline data storage.
- Whether the tool is custom, off-the-shelf or open source.
- Data protection and privacy.

This systematic review aims to identify RCTs that have used health-related data collection tools in LMICs and to evaluate the characteristics of the identified data collection tools according to the modified MST framework.

Similar reviews investigating data collection tools in LMICs have been conducted. A systematic review by Keating *et al* investigated electronic data collection tools used for outbreak response in the context of the COVID-19 pandemic.[18] The review identified 75 electronic data collection, analysis and management tools that were used during the pandemic.[18] It emphasised the importance of improving interoperability among different tools and software to effectively manage outbreaks in LMICs.[18] The review also highlighted the need for additional training on these tools and software.[18] Faruk *et al* conducted a review examining the screening tools used in LMICs to identify developmental delays encompassing a range from neurological to behavioural concerns in children.[19] A total of 16 tools were identified for qualitative synthesis.[19] The findings indicated a significant lack of culturally sensitive

tools in LMICs.[19] Furthermore, most of the tools failed to reach the expected specificity and sensitivity due to the lack of access to a gold-standard assessment tool.[19] However, there is yet to be a review conducted that examines data collection tools as a whole.

The objectives of this review are as follows:

- Categorise the types of health-related data collection tools currently being used in LMICs. This may include digital/manual, custom/off-the-shelf and wearable/non-wearable among others.
- Identify the primary differences in the attributes between the various health-related data collection tools used in RCTs in LMICs.
- Establish a robust framework (eg, modified MST) for researchers to assess the characteristics of health-related data collection tools.

REVIEW QUESTIONS

This systematic review aims to address the following questions:

- What are the health-related data collection tools that are used in RCTs in LMICs?
- What are the key differences in the attributes of health-related data collection tools that are used in RCTs in LMICs with the modified MST framework as a reference point?
- How suitable is the modified MST framework for healthcare researchers to evaluate the characteristics of health-related data collection tools in LMICs?

METHODS

Design

This protocol has been registered with the International Prospective Register of Systematic Reviews (PROSPERO) (registration number: CRD42023405738). It has been developed using the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA) checklist.

Inclusion and exclusion criteria

Inclusion criteria

- Published, peer-reviewed, RCTs that use a health-related data collection tool in an LMIC will be eligible for inclusion in the initial stage of the systematic review.
- Participants are adults aged 18+.
- Recruitment of participants exclusively from the local population.
- Publication dates between January 2005 and June 2023. The year 2005 is set as the publication year limit as mobile devices such as smartphones became available along with the roll-out of the internet to facilitate the transmission of data.
- English language-only articles.
- Published in full-text.

Exclusion criteria

- Quasi-randomised trials will be excluded.
- The RCT does not use a health-related data collection tool.
- Studies where participants are under 18 years of age.
- The study will be excluded if there are less than five MST characteristics addressed in the RCT.

Patient and public involvement

None.

SEARCH STRATEGY

A search of the literature will be conducted in four electronic databases: PubMed, CINAHL, Web of Science and Embase. The primary search strategy was designed for PubMed and adapted as appropriate for each of the databases. The full search strategy for all databases is presented in online supplemental appendix. The terms were slightly adapted to the search particulars (eg, truncations, wildcards (*)) and filters available for each database. The search was conducted in June 2023.

SELECTION PROCESS

Two review authors (RK and NA-S) will independently screen the titles and abstracts of search results based on the eligibility criteria. The full-text articles will be retrieved if the abstract contains limited information about the study. In the case where a relevant conference abstract is identified, RK and NA-S will contact the authors for the full-text article. Duplicate articles will be removed. Study authors will be contacted for clarification if eligibility is unclear. Disagreements will be resolved through discussion. A third author (JO'D) will adjudicate if the disagreement cannot be resolved. The list of the excluded studies and the reasons for their exclusion will be presented in a 'Characteristics of excluded studies' table. The study selection process will be outlined in a PRISMA flow diagram. Mendeley referencing software will be used to screen and determine the eligibility of all the references from the initial search.

DATA COLLECTION

Data from the included studies will be extracted by RK and NA-S independently into a data extraction table in Microsoft Excel. This includes the following (online supplemental file 1):

- Descriptive information about the study: DOI, author and year of publication, objective of the study, where the RCT was conducted, language and age of the study participants.
- Descriptive information about the data collection tool: type of data collection tool, online or offline data storage, whether the tool is custom, off the shelf or open source and data protection and privacy.
- Information relating to the characteristics of the MST framework: functionality, reliability, usability, efficiency, maintainability, portability, effectiveness,

cost–benefit, satisfaction, freedom from risk and context coverage.

ASSESSMENT OF RISK OF BIAS

Risk of bias will not be conducted for the selected RCTs, as this systematic review is focused on the data collection tools being used, not the outcome of the RCTs themselves. However, quasi experimental studies that do not involve randomisation will be excluded.

DATA SYNTHESIS

Data will be analysed using a narrative synthesis approach. The included studies and their outcomes will be presented in a table format and categorised based on the relevant parameters. Qualitative analysis will be used to assess the tools identified within each study against the criteria of the modified MST framework. There are 11 characteristics within the framework. These include functionality, reliability, usability, efficiency, maintainability, portability, effectiveness, cost–benefit, satisfaction, freedom from risk and context coverage. These characteristics are broken down into 32 subcharacteristics. For example, functionality involves assessing the suitability, accuracy, interoperability and security of a data collection tool. These characteristics and the subcharacteristics will be used to conduct a descriptive analysis of each selected data collection tool. Moreover, in future studies, researchers can use the same characteristics, in addition to the added ones, to evaluate the attributes of any tools they are investigating. Finally, the modified MST framework will be evaluated based on its efficacy of assessing the characteristics of a data collection tool used in an RCT.

ETHICS AND DISSEMINATION

Formal ethical approval is not required as primary data will not be collected in this study. The findings from this systematic review will be published in a peer-reviewed journal.

DISCUSSION

This systematic review is the first to focus on health-related data collection tools in RCTs conducted in LMICs. The review aims to provide a comprehensive and up-to-date assessment of the various data collection tools currently being used in RCTs in LMICs, categorise the type of data collection tools, and assess their characteristics and the challenges associated with deploying them in these settings.

Achieving sustained growth in health policy and systems research in developing countries is a systemic issue that requires significant reform to existing research laws and policies in high-income countries.[20] The implementation of facility-building measures, such as the adoption of advanced digital data collection methods, can play a crucial role in mitigating research capacity issues in LMICs.[20]

The findings of this review may have significant implications for researchers seeking to use data collection tools in LMICs. Researchers may be unaware of the available range of data collection tools leading them to develop a customised tool, which can be costly and time-consuming. By identifying and assessing the characteristics of the various data collection tools, this review will assist researchers in selecting an existing tool that will meet their research objectives.

It is important to note that most LMICs are non-English-speaking countries. Therefore, a limitation of this systematic review is the exclusion of potentially relevant non-English articles, as a result of the applied filters. The anticipated impact on the results is minimal, given the relatively small number of non-English articles available. Another limitation is that individual countries identified as LMICs were not included in the search string due to the massive volume of results generated. However, the search string has been modified to include different variations of the term ‘LMICs’ and the income classification of each country will be evaluated based on the World Atlas Bank’s definition.

Author affiliations
[1]University College Cork, Cork, Ireland
[2]School of Medicine, University College Cork, Cork, Ireland
[3]Epidemiology and Public Health, University College Cork, Cork, Ireland
[4]London School of Hygiene & Tropical Medicine, London, UK
[5]Primary Care and Public Health, Imperial College London, London, UK
[6]ASSERT Research Centre, University College Cork, Cork, Ireland
[7]Malawi eHealth Research Centre, University College Cork, Cork, Ireland

Twitter Frances Shiely @FrancesShiely, Titus Divala @TDivala and Kayode Philip Fadahunsi @PF_eHealth

Contributors RK, NA-S, JO'D, KPF and EJK designed the protocol. RK and NA-S wrote the first draft of the protocol. JO'D and KPF provided critical appraisal regarding the design of the systematic review and revised the protocol. RK and NA-S performed the search and designed the extraction sheet. All authors approved the final draft of the protocol. All authors meet the ICMJE criteria.

Funding The open access fee was paid from the Imperial College London Open Access Fund.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Rithvik Karthikeyan http://orcid.org/0009-0000-4539-518X
Noor Al-Shamaa http://orcid.org/0009-0006-1192-1253
Titus Divala http://orcid.org/0000-0003-3029-9579
Kayode Philip Fadahunsi http://orcid.org/0000-0003-1470-5493
John O'Donoghue http://orcid.org/0000-0001-6056-8640

REFERENCES

1 Kabir SMS. Methods of data collection [ResearchGate]. 2018. Available: https://www.researchgate.net/publication/325846997_METHODS_OF_DATA_COLLECTION
2 Tonyushkina K, Nichols JH. Glucose meters: a review of technical challenges to obtaining accurate results. *J Diabetes Sci Technol* 2009;3:971–80.
3 Alghamdi NS, Alghamdi SM. The role of digital technology in curbing COVID-19. *Int J Environ Res Public Health* 2022;19:8287.
4 Will DJ, Campbell KJ, Holmes ND. Using digital data collection tools to improve overall cost-efficiency and provide timely analysis for decision making during invasive species eradication campaigns. *Wildl Res* 2014;41:499.
5 Channa A, Popescu N, Skibinska J, *et al*. The rise of wearable devices during the COVID-19 pandemic: a systematic Review. *Sensors (Basel)* 2021;21:5787.
6 Rahman MA, Moayedikia A, Wiil UK. Editorial: data-driven technologies for future healthcare systems. *Front Med Technol* 2023;5:1183687.
7 Alshaikh F, Ramzan F, Rawaf S, *et al*. Social network sites as A mode to collect health data: a systematic review. *J Med Internet Res* 2014;16:e171.
8 Pratt B, Loff B. Health research systems: promoting health equity or economic competitiveness? *Bull World Health Organ*;90:55–62.
9 Katyal A, Kumar A. Strengthening emergency care by developing data collection systems in Low- and middle-income countries. Elsevier. 2020. Available: https://www.sciencedirect.com/science/article/pii/S2211419X20301002#s0025 [Accessed 6 Jun 2023].
10 World bank country and lending groups – world bank data help desk, Available: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups [Accessed 31 May 2023].
11 Akobeng AK. Understanding measures of treatment effect in clinical trials. *Arch Dis Child* 2005;90:54–6.
12 Rubagumya F, Hopman WM, Gyawali B, *et al*. Participation of lower and upper middle-income countries in cinical trials led by high-income Ccountries. *JAMA Netw Open* 2022;5:e2227252.
13 Fisher MB, Mann BH, Cronk RD, *et al*. Evaluating mobile survey tools (MSTs) for field-level monitoring and data collection: development of a novel evaluation framework, and application to MSTs for rural water and sanitation monitoring. *Int J Environ Res Public Health* 2016;13:840.
14 Connelly K, Molchan H, Bidanta R, *et al*. Evaluation framework for selecting wearable activity monitors for research. *Mhealth* 2021;7:6.
15 Stoyanov SR, Hides L, Kavanagh DJ, *et al*. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015;3:e27.
16 McMillan B, Hickey E, Patel MG, *et al*. Quality assessment of a sample of mobile app-based health behavior change interventions using a tool based on the National Institute of Health and Care excellence behavior change guidance. *Patient Educ Couns* 2016;99:429–35.
17 Zelmer J, van Hoof K, Notarianni M, *et al*. An assessment framework for E-mental health Apps in Canada: results of a modified delphi process. *JMIR Mhealth Uhealth* 2018;6:e10016.
18 Keating P, Murray J, Schenkel K, *et al*. Electronic data collection, management and analysis tools used for outbreak response in low- and middle-income countries: a systematic review and stakeholder survey. *BMC Public Health* 2021;21:1741.
19 Faruk T, King C, Muhit M, *et al*. Screening tools for early identification of children with developmental delay in low- and middle-income countries: a systematic review. *BMJ Open* 2020;10:e038182.
20 The Role of Europe and of International Science and Technology Cooperation. Council on health research for development - COHRED. Available: https://www.cohred.org/publications/open-archive/the-role-of-europe-and-of-international-science-and-technology-cooperation-specific-needs-for-health-research-in-low-and-middle-income-countries/ [Accessed 4 Jun 2023].