2018

Defining and Developing a Generic Framework for Monitoring Data Quality in Clinical Research

Lauren Houston
University of Wollongong, lah993@uowmail.edu.au

Ping Yu
University of Wollongong, ping@uow.edu.au

Allison Martin
University of Wollongong, allisonh@uow.edu.au

Yasmine Probst
University of Wollongong, yasmine@uow.edu.au

Defining and Developing a Generic Framework for Monitoring Data Quality in Clinical Research

Abstract

Evidence for the need for high data quality in clinical research is well established. The rigor of clinical research conclusions rely heavily on good quality data, which relies on good documentation practices. Little attention has been given to clear guidelines and definitions to monitor data quality. To address this, a "fit-for-use" data quality monitoring framework (DQMF) for clinical research was developed based on a holistic design-oriented approach. An integrated literature review and feasibility study underpinned the framework development. Ontology of key terms, concepts, methods, and standards were recorded using a consensus approach and mind mapping technique. The DQMF is presented as a nested concentric network illustrating concept relationships and hierarchy. Face validation was conducted, and common terminology and definitions are listed. The consolidated DQMF can be adapted according to study context and data availability aiding in the development of a long-term strategy with increased efficacy for clinical data quality monitoring.

Disciplines
Medicine and Health Sciences

Defining and Developing a Generic Framework for Monitoring Data Quality in Clinical Research

Miss Lauren Houston, BSc(Hons.)[1,2], A/Prof Ping Yu, PhD[3], Dr Allison Martin MSc, Grad Dip PH, PhD[1] and Dr Yasmine Probst MSc, MHlthInfo, GradCertBus, PhD, AdvAPD[1,2]

[1] School of Medicine, Faculty of Science, Medicine and Health, University of Wollongong, Wollongong NSW 2522, Australia; [2] Illawarra Health and Medical Research Institute, University of Wollongong, Wollongong NSW 2522, Australia; [3] Centre for IT-enabled Transformation, School of Computing and Information Technology, Faculty of Engineering and Information Sciences, University of Wollongong, Wollongong NSW 2522, Australia.

Abstract

Evidence for the need for high data quality in clinical research is well established. The rigor of clinical research conclusions rely heavily on good quality data, which relies on good documentation practices. Little attention has been given to clear guidelines and definitions to monitor data quality. To address this, a "fit-for-use" data quality monitoring framework (DQMF) for clinical research was developed based on a holistic design-oriented approach. An integrated literature review and feasibility study underpinned the framework development. Ontology of key terms, concepts, methods, and standards were recorded using a consensus approach and mind mapping technique. The DQMF is presented as a nested concentric network illustrating concept relationships and hierarchy. Face validation was conducted, and common terminology and definitions are listed. The consolidated DQMF can be adapted according to study context and data availability aiding in the development of a long-term strategy with increased efficacy for clinical data quality monitoring.

Introduction

Regardless of study design or clinical area, high quality data collection and standardized data processing and representation are paramount to ensure reliable research findings[1-3]. Evidence has linked poor data quality to incorrect conclusions and recommendations[4-7]. Preventing data error is key during the development, design, and collection of clinical data[8]. The National Institute of Health (NIH)[9] broadly defines a clinical trial as *"a research study in which one or more human subjects are prospectively assigned to one or more interventions (which may include placebo or other control) to evaluate the effects of those interventions on health-related biomedical or behavioral outcomes"* [para. 3]. With this broad definition, several challenges arise to ensure high data quality due to different clinical objectives and requirements for data. Many strategies and interventions have been developed and are aimed at reducing error in clinical trials, including standard operating procedures (SOPs), personnel training, data monitoring or auditing and careful design of case report forms (CRF). However, current international and national guidelines lack consistency, creating uncertainty for clinical researchers. Therefore, in order to optimize data quality, standard procedures must be implemented.

The need for high data quality in clinical research has been well established. The term 'data quality', sometimes referred to as information quality[10], is a multidimensional and hierarchical concept[11]. The World Health Organization (WHO)[12] defines data quality as " *...the ability to achieve desirable objectives using legitimate means. Quality data represents what was intended or defined by their official source, are objective, unbiased and comply with known standards*" [pg.10]. An alternate definition emerged from Kerr et al.[13] who suggested that good data quality is data "*fit for use*" [pg.5] for the objectives of data collection. High data quality is crucial to a research organization's success, while poor data quality (often referred to as 'dirty data') can significantly impact on the productivity of a business or institution[14]. It is essential to minimize errors, as a poorly designed study with inferior data points and results cannot be redeemed. In the context of clinical research, data that is not "fit for use" may lead to biased results, conclusions, and recommendations and may compromise participant health. To date, there is a lack of a precise definition for data quality. This, in turn, creates misunderstandings that may weaken the validity and reliability of data quality assessment and monitoring methods[4].

Data management needs to be consistent, effective and efficient within each study[8, 15]. Regardless of the method used to collect, handle and store the data collected within clinical trials, a vigorous management system is essential. For academic clinical trials, developing and maintaining a data management system is a challenge[16]. This is largely due to special

requirements for individual trials, for example, the need to implement specific frameworks, and the expense to develop and run the software which also requires a sophisticated information technology infrastructure[17]. Academic clinical trials are less likely to implement common clinical data management systems such as those utilized within the pharmaceutical industry (e.g. Oracle Clinical); instead they often implement specialized, in-house smaller systems[18]. A recent survey found that considerable heterogeneity in data management exists and limited open access or freely available standard documents are available[18]. Furthermore, over 50% of the surveyed clinical research centers stated that although they had a data management system in place, the system did not comply with guidelines and legal requirements (GCP, ECRIN, FDA, GAMP, and ISO) for both internal system and independent validation by an external auditor. Similarly, survey results from the Association of Academic Health Centers (AAHC) highlighted that the greatest barrier for clinical trial operations was a lack of resources, systems, and procedures within the organisation[19]. It is clear that clinical trials, especially within the academic research community, need a standardized, open-source data monitoring framework to improve the quality level and best practice.

With rapid developments in technology, clinical research now relies heavily on the evaluation of automatic and electronically communicated data for critical decision-making through which data quality has become increasingly important. Emerging literature on technological improvements, demonstrate new opportunities and concerns about the reuse of clinical research data. The American Medical Informatics Association have compiled recommendations and also stressed the urgency and complexity of issues that surround the secondary use of clinical data[20]. Additionally, the primary obstacle to integrated data repositories was data quality[21]. In an effort to overcome this issue and to continuously improve data quality; standard monitoring methods are required before, during and after primary data collection, and at a larger-scale for the reuse of clinical data in research. To optimize quality, clinical studies should implement and publish their approach to monitor data quality to increase efficacy, reduce costs and follow procedures designed to minimize inaccurate and incomplete data.

The aim of this research was to develop a "fit-for-use" data quality monitoring framework (DQMF) for clinical research. This framework will aid clinical trials in obtaining and maintain high data quality by providing guidance on critical areas that relate to trial operations throughout the clinical research process.

Framework development

When determining data quality criteria there are different approaches that can be applied, which include empirical, practitioner-based, theoretical, literature-based, pragmatic and design-oriented[22, 23]. For the purpose of this research, a holistic design-oriented approach was applied to design and develop the DQMF. A design-oriented approach provided guidance to the researchers to create the framework (artefact) and further understand the apparent reality of different stakeholders (clinical research trials) of the framework[24]. This provided further guidance to the researchers by helping to recognize data failures by developing the framework against a real-world state[25]. Design science is considered a problem-solving paradigm and seeks to extend the boundaries of the human and organizational capabilities by creating new and innovative artefacts[26]. The purpose of such artefacts is to improve the efficiency and effectiveness of the organization's characteristics, the work systems and its employee's capabilities. Design science argues that human knowledge and understanding of the problem and the solution are acquired in the 'building' and the 'application' of the artefact. Therefore, this research follows the conceptual framework and seven guidelines proposed by Hevner et al.[26] for understanding, executing and evaluating design-oriented research. This creative design process utilised a build-and-evaluate loop iteration to the evolving design of the generated artefact. This preliminary methodological research focuses largely on the 'build' process of the resulting artefact, keeping in mind that further evaluation and testing needs to be conducted. Overall, the proposed framework aims to help those designing, implementing and working in clinical research to understand the complex inter- and intra-relationships between the concepts that need to be planned both methodically and structurally, in order to improve the data quality of clinical research.

The initial design of the DQMF was guided by an integrated literature review and feasibility study[27] of data quality concepts, from an information sciences and clinical grounding. Outcomes from the feasibility study determined that clinical trials are implementing ad hoc methods pragmatically to ensure data quality. Thus, there is a necessity for further research into 'standard practice'. The ontology of key terms, concepts, methods, and standards were extracted and recorded from the literature review and feasibility study survey questions. Consensus approach and mind mapping techniques were used to present associations in a non-linear diagram/network[28]. The dependent variable, 'data quality monitoring' was placed at the centre of the network to compose the mind map where associations were added and 'branched'. This process was undertaken by the researchers (L.H., P.Y., A.M. and Y.P.) to foster a natural thinking process, allowing for the addition of new concepts,

relationships and annotations[29]. Furthermore, branches and nodes were grouped together under comparable topic areas via researcher agreement to construct a hierarchical tree-like figure.

Once the researchers (L.H., P.Y., A.M. and Y.P.) came to a consensus, the DQMF was presented to a convenient sample (n=8) of working health professionals (dieticians, nutritionists, educators, public health practitioners) for face validity testing. Participants had clinical research experience (1 – 15+ years) in university academic, private institute and hospital settings. The primary researcher (L.H) moderated the interactive one-hour workshop, which aimed to gain feedback on the design and useability of the proposed DQMF within different clinical research settings. Each participant was provided an individual copy of the DQMF and encouraged to make note of any questions or issues. The workshop communicated the process by which the artefact was created and defined as the mechanism to finding an effective and efficient solution. Once the background information was presented, participants were asked to refine and make relevant changes to the DQMF based on their own knowledge and expertise. The primary researcher (L.H) then opened up the discussion to the group to explore reasons why amendments were suggested to fit each of the participant's clinical focus. The workshop identified that standardized terminologies, definitions and dialogue are crucial to the success of the DQMF. According to participants' responses, amendments to the DQMF were discussed and agreed upon and a supporting list of key terms and definitions were devised. Each stage of the systematic framework development was aligned with the international guidelines (GCP, ECRIN, FDA, GAMP, and ISO) to ensure the taxonomy and terminology used complied with global standard procedures and policies.

Data Quality Monitoring Framework (DQMF)

Refinement and evaluation of the key concepts has led to the development of the DQMF. This framework contains the key components of data quality, data quality monitoring, and data quality management presented in a nested concentric network to illustrate the relationships and hierarchy (Figure 1). Each layer of the framework contains specific and highly important procedures and concepts undertaken within each layer. The importance of training and education is highlighted by its expansion across all layers of the DQMF. It was determined in the stakeholder workshop that dialogue, definitions, and terminology should be implemented consistently across clinical research. Due to the need to clarify terminology related to the DQMF and ensure effective communication we have included Table 1, which highlights key terms and their definitions related specifically to the DQMF.
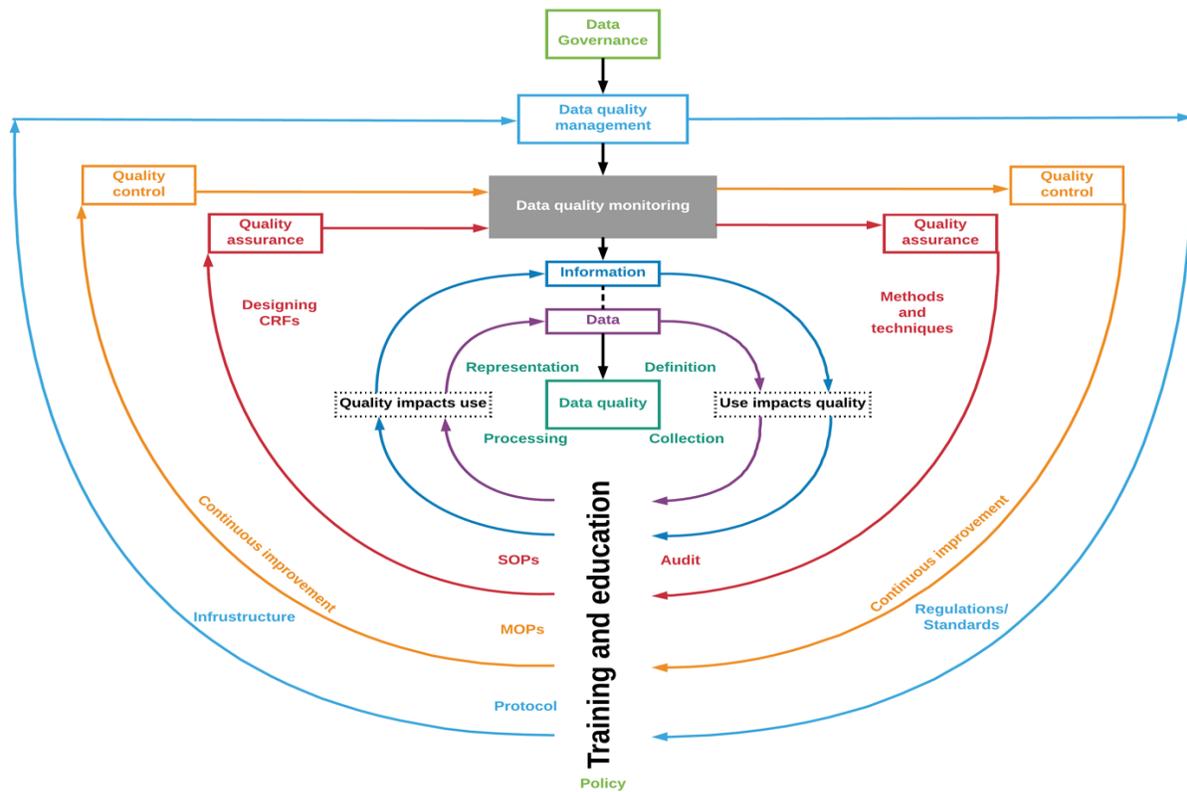
Figure 1: The data quality monitoring framework

Four key independent variables were identified (inner layer) and adapted from Nahm[30] who illustrates the way data definition, collection, processing and representation are each handled, impacts data use which impacts on data and information quality. On the contrary, data and information quality in turn impact use. The data evolution life cycle (DELC) was deliberated for inclusion within the inner tier throughout the development of the DQMF as it reflects a sequence of stages known as data collection, organization, presentation, and application[31]. The researchers chose not to integrate this cycle as they believe the stages of collection, organization and presentation relate closely to the terminology and stages of Nahm's[30] framework of collection, processing and representation, respectively. Additionally, the DQMF main focus is 'data quality monitoring' in which Nahm's framework was designed to highlight the factors that affect data quality while the DELC represents data evolution. The addition of 'application' was discussed, however, the researchers believe that the framework illustrates how data is utilized and applied in clinical research by linking 'data' to 'information'.

Data quality monitoring was separated into two main concepts; quality assurance and quality control (middle tiers). The terms quality assurance and quality control are often used inaccurately or interchangeably[8]. Quality assurance is the process to "prevent" data errors, which includes the methods such as audits and other methods/techniques to ensure data integrity. Auditing is a recognized method that has been used to assess and develop the quality of information[32, 33]. Quality assurance audits within clinical and healthcare settings are extensively employed and are the major strategy to ensure high-quality data[30, 34-36]. Quality assurance activities include examining the design of case report forms, analyzing the data collection techniques and regular training of data entry personnel and data management[37, 38]. On the other hand, quality control is the process to "alleviate" or "remove" the impact of errors that have occurred during data collection and/or analysis[15]. This refers to the operational techniques used to fulfil requirements for quality. Recognised methods of quality control include the conduct of periodic monitoring (daily, weekly, and monthly) through pragmatic data range and consistency checks, query management, and double data entry to minimize errors[8]. Therefore, quality control is the continuous quality assurance activity undertaken to verify clinical trial-related processes to fulfil the agreed standards.

Data quality management and data governance (the outer tiers) include developing and implementing national and internal standards and regulations in the full data life cycle, including planning before and execution of protocols and policies for data capture and analysis. Currently, there is no open access Good Clinical Practice (GCP) standard on data monitoring that is broadly recognized, detailed and applicable to all clinical research. Therefore, the design of the DQMF has drawn upon an illustration by Ohmann et al.[16] who highlight the central importance of the International Conference on Harmonisation (ICH)-GCP guidelines within clinical research at an international level and within the United States of America regulations and the European Union directives. The illustration links regulations and guidelines that are relevant to GCP-compliant computer systems and data management practices connecting important references from one document to another. The researchers took guidance from Ohmann's work and extracted key concepts within each of the regulatory requirements and documents. In light, the researchers agreed a simpler and broader approach was required to guide clinical researchers by providing overarching concepts of infrastructure, protocol and regulations/standards. This approach will allow clinical researchers to make an informed decision regarding the most suitable management strategy to their individual trial while incorporating and highlighting legal requirements. Further, it will acknowledge the broad range of clinical research trials and the fact that healthcare requires person-to-person interactions for collaboration and integration between strategy, process (automated/non-automated) and the supporting information systems. As the feasibility study that guided this research found heterogeneity in data management practices with only 50% of respondents reporting to have a data management plan in place[27]. By providing a consolidated framework to optimize the efficacy of data management, the researchers aim to provide clear guidance to clinical research data quality monitoring that is both time and cost effective.

Table 1: DQMF terms, abbreviations, and definitions

Term/ Abbreviation	Definition
Audit	A systematic and independent examination of trial-related activities and documents to determine whether these activities were conducted, and the data were recorded, analyzed and accurately reported according to the protocol, the sponsor, SOP, GCP and applicable regulatory requirements[39].
Compliance	Adherence to all the trial-related requirements, GCP requirements and applicable regulatory requirements[39].

Continuous improvement	Systematically planning, collecting and assessing data to distinguish a chance to ameliorate the process of clinical trial data management[40, 41].
CRF/eCRF	Case report form/electronic case report form: A printed, optical or electronic document designed to record all of the protocol-required information[39].
Data	An individual fact or pieces of information.
Data collection	The process by which data elements are accumulated[3].
Data definition	Occurs as the protocol or research plan is developed. Procedures include identifying data to be collected, defining data elements and designing CRFs[3].
Data governance	Encompasses people, processes and information technology required to formally manage and exercise control over methods used by data stewards and data custodians in order to improve data quality[42].
Data processing	The processes and systems applied to audit and monitor data[3].
Data representation	Relates to data analysis and the process of translating data into meaningful information[3].
Data quality	Data quality can be defined as the degree to which a set of characteristics of data fulfills requirements[43]
Data quality management	The development, execution, and supervision of plans, policies, programs and practices that control, protect, deliver and enhance the value of data and information assets[44].
Data quality monitoring	The oversight and review of research processes, procedures, records, data reporting, appropriate conduct and ongoing evaluation.
Education and training	The knowledge or skills obtained or developed by a learning process and the further instruction and education to an agreed standard of proficiency.
Information	A collection of data or facts.
Infrastructure	Buildings, supplies, policies, procedures, information technology and other assets that support the human resources of an organization.
Methods and techniques	*Risk-based monitoring:* A mixed method approach focused on the critical data points and processes that are identified to have the most risk via a targeted or triggered assessment[27].
	Remote monitoring: Data monitored off-site, includes delivering documents via email, fax or snail mail to monitoring personnel to conduct source data verification[27].
	Centralised monitoring: Data collected through an electronic data capture and queries identified by a monitor that may need further attention to alleviate problems[27].
	Source data verification: Comparing source data (original or certified copy) documents to data recorded or entered into a case report form or electronic record or database[27].
	On-site monitoring: All monitoring activities undertaken at the clinical trial site[45].
Manual of operations (MOP)	A handbook of instructions designed to guide the research team to successfully carry out aspects of a research project according to the research protocol[46].
Monitoring	The act of overseeing the progress of a clinical trial, and of ensuring that it is conducted, recorded, and reported in accordance with the protocol, SOPs, GCP and the applicable regulatory requirement(s)[39].
Policy	Communicates and documents an organizations overall intentions and direction with respect to quality. A written quality policy, and top-level management should demonstrate commitment to the quality policy by supporting the organization's infrastructure with adequate

	resources. Off-line quality control activities, such as quality engineering, quality planning, and procedures applicable to each study, will be enhanced by this infrastructure and facilitate error prevention[42].
Protocol	A document that describes the objective(s), design, methodology, statistical considerations and organization of a trial. The protocol usually gives the background and rationale for the trial, but these could be provided in other protocol referenced documents[39].
Quality assurance	All those planned and systematic actions that are established to ensure that the trial is performed, and the data are generated, documented (recorded), and reported I compliance with GCP and applicable regulatory requirements[39].
Quality control	The operational techniques and activities undertaken within the quality assurance system to verify that the requirements for quality of the trial-related activities have been fulfilled[39].
Regulations/ standards	*International Organisations for Standardisation (ISO)[47]:* ISO 14155:2011 – Clinical investigation of medical devices for human subjects – Good Clinical Practice: This International Standard specifies general requirements intended to; protect the rights, safety and well-being of human subjects, ensure the scientific conduct of the clinical investigation and the credibility of the clinical investigation results, define the responsibilities of the sponsor and principal investigator, and assist sponsors, investigators, ethics committees, regulatory authorities and other bodies involved in the conformity assessment of medical devices.
	International Conference of Harmonisation (ICH) - Good Clinical Practice (GCP) E6(R2)[39]: An international ethical and scientific quality standard for designing, conducting, recording and reporting trials that involved the participation of human subjects.
	*Regulatory authorities: Bo*dies having the power to regulate. Individual countries different regulatory requirements and enforcement abilities, for example, The United States (US) Food and Drug Administration (FDA) Guidelines for Monitoring of Clinical Investigations[45], the European Union's (EU) Clinical Trial Directive[48] and the Australian Government's National Statement on Ethical Conduct in Human Research[49].
Standard Operating Procedure (SOP)	Detailed written instructions to achieve uniformity of a specific function. A written process of instructions necessary to carry out a policy or a way a task can be performed the same way each time[39].

Discussion

A systematic approach to data quality monitoring is essential to ensure high data quality for confidence in data reuse and technological improvements for clinical research. The DQMF developed provides an easy-to-use guide for monitoring data quality of individual clinical research trials with reference to key international documents. Within the pharmaceutical/private industries [40] and information sciences literature, data quality audit tools and procedures appear to be well developed with many frameworks acknowledging the multiple dimensions of data quality [50-55]. However, only a small body of clinical research has described the use of data quality frameworks [22, 40, 56-58], and even less have identified the appropriate methods to quantify the quality of data [30]. Although many data quality dimensions and attributes have been determined within the clinical and health literature, the majority provide no usable definitions. Public sharing of such knowledge is crucial in developing a standardized approach that can be implemented across the clinical and broader research community to improve the rigor of clinical research.

Many organizations collect and analyze data for their own benefit to meet SOPs and ensure quality assurance and control. In terms of quality assurance and quality control the SOP is one of the most generic, reusable and important documents within clinical research[15]. However, within academic research settings, published on-site audits that include quality assurance are less often reported. This may be due to unclear audit methods, lack of time and funding, audits perceived as unnecessary for unregulated studies and publishing SOPs is not seen as a 'value added' activity[59-61]. There is a general agreement among

leading clinical trial management groups that establishing reliable guidelines as a monitoring strategy would need to be determined on a risk-adapted basis for each trial[36]. It is recognized that different strategies need to be tailored for different types of clinical trials to determine adequate and appropriate monitoring[41, 62]. However, published methodology papers are warranted to promote routine auditing and monitoring within both academic and commercial research settings. Research grants seldom include funding for such programs[63]. The DQMF has gathered current published information on the conduct of clinical trial data management, albeit limited. The application of the framework is a vital implementation strategy to the overall improvement of the quality of clinical trial data and the follow-on-effect of results, conclusions and recommendations[60]. Identifying all possible data discrepancies before they occur with all best intentions may not happen; therefore, a standardized framework, such as the one in this study, will provide useful guidance for the pragmatic implementation of continuous quality improvement.

The interest in standardization within the clinical research community has grown in recent years and therefore, the DQMF considers data quality monitoring from a broad perspective. This generic framework brings together key concepts from the scientific literature, government documents, and policies to illustrate links between concepts and their effect on each other within and amongst layers. This differs from previously published frameworks, which have focused on specific concepts in isolation, not considering the inter- and intra-relationships. This singular approach has caused confusion within the clinical research space. A survey conducted by the Clinical Trials Transformative Initiative (CTTI) found heterogeneity in data quality monitoring intensity, focus and methodology within and between academic/government, industry and clinical research organisations[64]. The utility of the DQMF is that it provides a single consolidated framework, which allows adaptations according to study context and data availability. This research will benefit the development of a long-term strategy and focus to fill the knowledge gap and reduce confusion around data quality monitoring in clinical research trials. Currently, as no standardized definitions apply across all clinical context. In correspondence with an increasing movement from paper-based forms to a digital and adaptive learning environment, it is necessary to improve the methods and approach to collection, storage and sharing of clinical research data. Electronic solutions are relatively new in clinical research and require major changes to existing procedures and professional training. Additionally, challenges arise in incorporating electronic data standards (CDISC and HL7) and the role they play in ensuring efficient and economic data sharing within clinical research[65]. The proposed DQMF provides guidance to clinical researchers on areas related to trial operations and ensuring high data quality throughout the entire research process. By utilizing this generic framework, it is anticipated to minimize the obstacles related to primary data quality and for the reuse of clinical research data. As the DQMF continues to evolve throughout the design-orientated approach, our knowledge and understanding of the challenges that arise from adapting to an electronic world will be addressed. This is vital to ensure the generic framework has future applicability.

The holistic designed-orientated approach provided guidance to developing the DQMF by aiding the researchers to understand the clinical stakeholders. The framework aims to improve clinical research trial practices, which currently consists of complex, isolated and independent tools, procedures and frameworks. By providing an easily integrated knowledge development tool for clinical research practice the DQMF will support clinical research as a value-added function by providing oversight and guidance on the complex area of data quality monitoring. Additionally, providing clear definitions to concepts are key to its success. It should be highlighted that the proposed DQMF is developed from the published literature and draws on the personal experiences of the research team and the participants included in the face validation workshop. This may be considered a bias. A major limitation of the proposed DQMF is that it is yet to be applied in practice and implemented in a clinical research trial. The researchers stress the importance that application of the DQMF is required to test the framework within a broad spectrum of clinical research studies to identify facilitators and barriers, thereby ensuring best practice for data quality. According to the design-science research guidelines, further evaluation, contribution, rigor, and communication are needed to develop a convincing argument for the utility of this framework and its purpose for real world use. It is suggested that empirical research be conducted through the use of a reactive Delphi study[66] to validate and allow experts to reach a consensus of opinion on the illustration, terms and what constitutes data quality monitoring in clinical research. The importance of industry wide definitions and methods are essential to enable strategic management and evaluate quality information. Without standardization, principle investigators of clinical research are left with inefficient data quality management.

Conclusion

A data quality monitoring framework (DQMF) has been developed for clinical research trials. The utility of this single consolidated framework is to reduce confusion around data quality monitoring whilst allowing for adaptations according to study context and data availability. The framework will guide new trials or identify procedures in existing trials to improve

data quality monitoring. The framework demonstrates how data quality monitoring develops over the life cycle of a clinical study and how knowledge management may guide new approaches to research. The DQMF must now be validated by applying the framework and terminology to various clinical research trials for real world use. This will be crucial to refine and evaluate the generic detailed framework. Overall, the DQMF will aid in the development of a long-term strategy to increase efficacy for clinical research data quality monitoring.

List of Abbreviations

AAHC: Association of Academic Health Centers
CDISC: Clinical Data Interchange Standards Consortium
CRF: Case report form
CTTI: Clinical trials Transformative Initiative
DELC: Data evolution life cycle
DQMF: Data quality monitoring framework
ECRIN: European Clinical Research Infrastructure Network
EU: European Union
FDA: Food and Drug Administration
GAMP: Good Automated Manufacturing Practice
GCP: Good Clinical Practice
HL7: Health Level Seven
ICH: International Conference on Harmonisation
ISO: International Organisation for Standardisation;
MOP: Manual of operations
NIH: National Institute of Health
SOP: Standard Operating Procedure
US: United States
WHO: World Health Organisation

References

1. Bhatt A. Quality of Clinical Trials: A Moving Target. Perspectives in Clinical Research. 2011;2(4):124-128.
2. Bowman S. Impact of Electronic Health Record Systems on Information Integrity: Quality and Safety Implications. Perspectives in Health Information Management. 2013(10Fall):1c.
3. Nahm M. Clinical Research Informatics In: Richesson RL, Andrews JE, editors. Health Informatics. London: Springer; 2012. p. 175-201.
4. Chen H, Hailey D, Wang N, Yu P. A Review of Data Quality Assessment Methods for Public Health Information Systems. International Journal Of Environmental Research And Public Health. 2014;11(5):5170-5207.
5. Hasan S, Padman R. Analyzing the Effect of Data Quality on the Accuracy of Clinical Decision Support Systems: A Computer Simulation Approach. AMIA Annual Symposium Proceedings 2006:324-328.
6. Huneycutt BJ, Illes J, Boone C, Jackson K, Woolett G, Health A. In the Patient's Interest: Improving Access to Clinical Trial Data. 2014. Available from: http://www.fdanews.com/ext/resources/files/09/09-03-2014-Avalerereport.pdf.
7. Kahn MG, Brown JS, Chun AT, Davidson BN, Meeker D, Ryan PB, et al. Transparent Reporting of Data Quality in Distributed Data Networks. EGEMS (Washington, DC). 2015;3(1):1052-1052.
8. Society for Clinical Data Management Good Clinical Data Management Practices. 2013.
9. National Institute of Health (NIH), NIH Clinical Trials Definition. 2014. Department of Health & Human Services 2014. United States. Available from: https://www.nidcd.nih.gov/research/clinical-studies/researchers-professionals/clinical-trials-definition.
10. Ge M, Helfert M. A Review of Information Quality Research - Develop a Research Agenda. Proceedings of the 12th International Conference on Information Quality. 2007:76--91.
11. Wang RY, Reddy MP, Kon HB. Toward Quality Data: An Attribute-Based Approach. Decision Support Systems. 1995;13(3):349-372.
12. World Health Organisation (WHO), Improving Data Quality a Guide for Developing Countires. 2003. WHO Regional Office for the Western Pacific. 2003. Switzerland. Available from: http://www.wpro.who.int/publications/docs/Improving_Data_Quality.pdf.

13. Kerr KA, Norris N, Stockdale R. The Strategic Management of Data Quality in Healthcare. Health Informatics Journal. 2008;14(4):259-266.

14. Haug A, Zachariassen F, van Liempd D. The Costs of Poor Data Quality. Journal of Industrial Engineering and Management. 2011;4(2):168.

15. Gutierrez JB, Harb OS, Zheng J, Tisch DJ, Charlebois ED, Stoeckert CJ, Jr., et al. A Framework for Global Collaborative Data Management for Malaria Research. The American Journal Of Tropical Medicine And Hygiene. 2015;93(3 Suppl):124-132.

16. Ohmann C, Kuchinke W, Canham S, Lauritsen J, Salas N, Schade-Brittinger C, et al. Standard Requirements for GCP-Compliant Data Management in Multinational Clinical Trials. Trials. 2011;12(1):85-85.

17. Krishnankutty B, Shantala B, Kumar BRN, Moodahadu LS. Data Management in Clinical Research: An Overview. Indian Journal of Pharmacology. 2012;44(2):168-172.

18. Kuchinke W, Ohmann C, Yang Q, Salas N, Lauritsen J, Gueyffier F, et al. Heterogeneity Prevails: The State of Clinical Trial Data Management in Europe - Results of a Survey of ECRIN Centres. Trials. 2010(1):79.

19. Rubin E, Lazar D. Clinical Trials Offices: What's New in Research Administration. Association of Academic Health Centers. 2009.

20. Safran C. Toward a National Framework for the Secondary Use of Health Data: An American Medical Informatics Association White Paper. J Am Med Informatics Assoc. 2007;14(1):1-9.

21. MacKenzie SL, Wyatt MC, Schuff R, Tenenbaum JD, Anderson N. Practices and Perspectives on Building Integrated Data Repositories: Results from a 2010 Ctsa Survey. J Am Med Inform Assoc. 2012;19(e1):e119-124.

22. Mettler T, Rohner P, Baacke L. Improving Data Quality of Health Information Systems: A Holistic Design-Oriented Approach. European Conference on Information Systems (ECIS) 2008 proceedings Paper 164.

23. Patsopoulos NA. A Pragmatic View on Pragmatic Trials. Dialogues in Clinical Neuroscience. 2011;13(2):217-224.

24. Wang RY, Kon HB, Madnick SE, editors. Data Quality Requirements Analysis and Modeling. the Ninth International Conference on Data Engineering 1993; Vienna, Austria.

25. Leitheiser RL, editor Data Quality in Healthcare Data Warehouse Environments. the 34th Hawaii International Conference on Systems Science 2001; Hawaii, USA.

26. Hevner AR, March ST, Park J, Ram S. Design Science in Information Systems Research. MIS Quarterly. 2004;28(1):75-105.

27. Houston L, Probst Y, Yu P, Martin A. Exploring Data Quality Management within Clinical Trials. Applied clinical informatics. 2018;9(1):072-081.

28. Dhindsa HS, Makarimi K, Roger Anderson O. Constructivist-Visual Mind Map Teaching Approach and the Quality of Students' Cognitive Structures. Journal of Science Education and Technology. 2011;20(2):186-200.

29. Meier PS. Mind-Mapping. Social Research UP-DATE. 2007(52):1-4.

30. Nahm ML, Pieper CF, Cunningham MM. Quantifying Data Quality for Clinical Trials Using Electronic Data Capture. PLoS ONE. 2008;3(8):1-8.

31. Liu L, Chi LN. Evolutional Data Quality: A Theory-Specific View. Proceedings of the Seventh International Conference on Information Quality (ICQ-02). 2002:292–304.

32. Kuperman GJ, Blair JS, Franck RA, Devaraj S, Low AF. Developing Data Content Specifications for the Nationwide Health Information Network Trial Implementations. J Am Med Informatics Assoc. 2010;17(1):6-12.

33. Duda SN, Wehbe FH, Gadd CS. Desiderata for a Computer-Assisted Audit Tool for Clinical Data Source Verification Audits. Studies In Health Technology And Informatics. 2010;160(Pt 2):894-898.

34. Mealer M, Moss M, Kittelson J, Sokol RJ, Kahn MG, Thompson BT, et al. Remote Source Document Verification in Two National Clinical Trials Networks: A Pilot Study. PLoS ONE. 2013;8(12):e81890.

35. Rostami R, Nahm M, Pieper CF. What Can We Learn from a Decade of Database Audits? The Duke Clinical Research Institute Experience, 1997--2006. Clinical Trials (London, England). 2009;6(2):141-150.

36. Macefield RC, Beswick AD, Blazeby JM, Lane JA. A Systematic Review of on-Site Monitoring Methods for Health-Care Randomised Controlled Trials. Clin Trials. 2013;10(1):104-124.

37. Clarke D, Breen L, Jacobs M, Franklin R, Tobota Z, Maruszewski B, et al. Verification of Data in Congenital Cardiac Surgery. Cardiology in the Young. 2008;18(S2):177-187.

38. Whitney CW, Lind BK, Wahl PW. Quality Assurance and Quality Control in Longitudinal Studies. Epidemiol Rev. 1998;20(1):71-80.

39. International Conference on Harmonisation (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use 2015, E6(R2) Good Clinical Practice (GCP) 2015. Available from: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM464506.pdf.

40. Arts DGT, De Keizer NF, Scheffer GJ. Defining and Improving Data Quality in Medical Registries: A Literature Review, Case Study, and Generic Framework. J Am Med Informatics Assoc. 2002;9(6):600-611.

41. Knatterud GL, Rockhold FW, George SL, Barton FB, Davis CE, Fairweather WR, et al. Guidelines for Quality Assurance in Multicenter Trials: A Position Paper. Controlled Clinical Trials. 1998;19(5):477-493.

42. Rosenbaum S. Data Governance and Stewardship: Designing Data Stewardship Entities and Advancing Data Access. Health Services Research. 2010;45(5 Pt 2):1442-1455.

43. International Standards Organisation (ISO), Quality Management Systems -- Fundamentals and Vocabulary Iso 9000:2015. 2015. Available from: https://www.iso.org/standard/45481.html.

44. Mosley M, Brackett MH, International D, Earley S, Henderson D. The Dama Guide to the Data Management Body of Knowledge: (Dama-Dmbok Guide): Technics Publications; 2010.

45. Food and Drug Adminisatration (FDA). FDA, Guidance for Industry, Oversight of Clinical Investigations - a Risk-Based Approach to Monitoring 2013. Available from: http://www.fda.gov/downloads/Drugs/.../Guidances/UCM269919.pdf.

46. Nation Institute of Health (NIH), Guidelines for Developing a Manual of Operations and Procedures (MOP), National Center for Complementary and Integrative Health (NCCIH). 2012. Available from: https://nccih.nih.gov/sites/nccam.nih.gov/.../MOP_NCCIH_ver1_07-17-2015.docx.

47. International Standards Organisation (ISO), Clinical Investigation of Medical Devices for Human Subjects - Good Clinical Practice ISO14155:2011. 2011. Available from: https://www.iso.org/obp/ui/#iso:std:iso:14155:ed-2:v1:en.

48. European Commisson, Clinical Trials - Directive 2001/20/EC 2001. Public Health Medical Products for Human Use. Available from: https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-1/dir_2001_20/dir_2001_20_en.pdf.

49. Australian Government, National Health and Medical Research Council (NHMRC), Australia Clinical Trials, Trial Sites. 2016. Department of Industry Innovation and Science. Available from: https://www.australianclinicaltrials.gov.au/trial-websites.

50. Scannapieco M, Catarci T. Data Quality under the Computer Science Perspective. Ital J Arch Comput. 2002;2:1-12.

51. Batini C, Cappiello C, Francalanci C, Maurino A. Methodologies for Data Quality Assessment and Improvement. ACM computing surveys (CSUR). 2009;41(3):16.

52. Wang RY, Storey VC, Firth CP. A Framework for Analysis of Data Quality Research. IEEE transactions on knowledge and data engineering. 1995;7(4):623-640.

53. Tayi GK, Ballou DP. Examining Data Quality. Communications of the ACM. 1998;41(2):54-57.

54. Pipino LL, Lee YW, Wang RY. Data Quality Assessment. Communications of the ACM. 2002;45(4):211-218.

55. Maydanchik A. Data Quality Assessment: Technics publications; 2007.

56. Davidson B, Lee YW, Wang R. Developing Data Production Maps: Meeting Patient Discharge Data Submission Requirements. Int J Healthcare Technology and Management. 2004;6(2):223-240.

57. Riain C, Helfert M. An Evaluation of Data Quallity Related Problem Patterns in Healthcare Information Systems. IADIS virtual multi conference on computer science and information systems 2005; Lisbon, Portugal.

58. Kahn MG, Raebel MA, Glanz JM, Steiner JF, Riedlinger K. A Pragmatic Framework for Single-Site and Multisite Data Quality Assessment in Electronic Health Record-Based Clinical Research. Medical Care. 2012;50(SUPPL. 1):S21-S29.

59. Califf RM, Karnash SL, Woodlief LH. Developing Systems for Cost-Effective Auditing of Clinical Trials. Controlled Clinical Trials. 1997;18(6):651-660.

60. Gardner KL, Dowden M, Togni S, Bailie R. Understanding Uptake of Continuous Quality Improvement in Indigenous Primary Health Care: Lessons from a Multi-Site Case Study of the Audit and Best Practice for Chronic Disease Project. Implementation Scince. 2010;5(1):21.

61. Mphatswe W, Mate KS, Bennett B, Ngidi H, Reddy J, Barker PM, et al. Improving Public Health Information: A Data Quality Intervention in Kwazulu-Natal, South Africa. Bulletin of the World Health Organization. 2012;90(3):176-182.

62. Prescott RJ, Counsell CE, Gillespie WJ, Grant AM, Russell IT, Kiauka S, et al. Factors That Limit the Quality, Number and Progress of Randomised Controlled Trials. Health Technology Assessment. 1999;3(20):1-143.

63. Ward R. Examining Methods and Practices of Source Data Verification in Canadian Critical Care Randomised Controlled Trials. ProQuest, UMI Disseration Publishing2013.

64. Morrison BW, Cochran CJ, White JG, Harley J, Kleppinger CF, Liu A, et al. Monitoring the Quality of Conduct of Clinical Trials: A Survey of Current Practices. Clinical Trials (London, England). 2011;8(3):342-349.

65. Gupta SK. Paperless Clinical Trials: Myth or Reality? Indian Journal of Pharmacology. 2015;47(4):349-353.

66. Ash S, O'Connor J, Anderson S, Ridgewell E, Clarke L. A Mixed-Methods Research Approach to the Review of Competency Standards for Orthotist/Prosthetists in Australia. International Journal of Evidence-Based Healthcare. 2015;13(2):93-103.