Data Quality Driven Improvement in Healthcare: A systematic Literature Review.

Anthony Lighterness, Michael Thomas Adcock, Lauren Abigail Scanlon, Gareth Price

Submitted to: Journal of Medical Internet Research
on: March 11, 2024

Table of Contents

Data Quality Driven Improvement in Healthcare: A systematic Literature Review.

Anthony Lighterness[1] BBiomedSc, MPhil, MSc; Michael Thomas Adcock[1] BSc, MSc; Lauren Abigail Scanlon[1] PhD; Gareth Price[2] PhD

[1]Clinical Outcomes and Data Unit The Christie NHS Foundation Trust Manchester GB
[2]Radiotherapy Related Research Group University of Manchester Manchester GB

Corresponding Author:
Anthony Lighterness BBiomedSc, MPhil, MSc
Clinical Outcomes and Data Unit
The Christie NHS Foundation Trust
550 Wilmslow Road
Manchester
GB

Abstract

Background: The promise of real-world evidence and the learning healthcare system primarily depends on access to high-quality data. Despite widespread awareness of the prevalence and potential impacts of poor data quality (DQ), current best practices for its assessment and improvement are unknown.

Objective: We sought to investigate how studies define, assess, and improve the quality of structured real-world healthcare data.

Methods: A systematic literature search of studies in the English language was implemented in EMBASE and PubMed databases to select studies that specifically aimed to measure and improve the quality of structured real-world data within any clinical setting. The time frame for the analysis was from January 1945 to June 2023. We standardised DQ concepts according to the DAMA (Data Management Association) DQ framework to enable comparison between studies. After screening and filtering by two independent authors, we identified 39 relevant articles reporting DQ improvement initiatives.

Results: Studies were characterised by considerable heterogeneity in settings and approaches to DQ assessment and improvement. Affiliated institutions were from 18 different countries and 18 different health domains, and most targeted data generated by multiple institutions. DQ assessment methods were largely manual and targeted only completeness and/or one other DQ dimension. Use of DQ frameworks (n = 6) and quality improvement methodologies (n = 5) were sorely lacking. Most studies reported improvements in DQ through the implementation of a combination of interventions, which included either DQ reporting and personalised feedback, implementation of new digital systems for electronic data capture, training on DQ or data standards, or improvements in clinical workflows. Reporting of changes in DQ varied significantly, making it difficult to conduct objective meta-analysis for determination of treatment effect.

Conclusions: There is an urgent need for standardised guidelines in the context of DQ improvement research to enable comparison and effective synthesis of lessons learnt. Frameworks such as Plan-Do-Study-Act (PDSA) learning cycles and the DAMA DQ framework can facilitate this unmet need. However, DQ improvement studies also need to pay closer attention to root cause analysis of DQ issues to ensure the most appropriate intervention is implemented, thereby ensuring long term sustainable improvement.

(JMIR Preprints 11/03/2024:57615)
DOI: https://doi.org/10.2196/preprints.57615

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✔ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✔ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Data Quality Driven Improvement in Healthcare: A systematic Literature Review

Authors: Anthony Lighterness[1], Michael Adcock[1], Lauren Scanlon[1], Gareth Price[1,2]

[1]The Christie NHS Foundation Trust, Manchester M20 4BX, UK
[2]University of Manchester, Manchester, UK

Corresponding Author:
Anthony Lighterness
Email: anthony.lighterness@nhs.net

Introduction

Background

The landscape of healthcare, improvement science, and digital technologies increasingly hinges on real-world data (RWD) to improve patient care and outcomes [1,2]. RWD encompasses a vast and dynamic collection of health-related information generated by means of routine clinical care from a diverse range of sources, such as electronic health records (EHRs), electronic medical records (EMRs), hospital information systems (HIS), picture archiving and communication systems (PACS), national registries, claims data, and wearable devices [2–4]. Despite their long history, the adoption and use of EHRs have become widespread only during the last decade [4,5]. EHRs and EMRs are often used interchangeably in the literature discussing health RWD, where some suggest that EMRs are a subset of EHRs [6], but the prominence of EHRs have positioned them as a primary source of RWD due to the comprehensive spectrum of patient information covered from genetic testing to treatment modalities and clinical outcomes [7,8]. To reflect this primary focus, the term “EHR-RWD” will be used throughout this review to denote real-world data derived from electronic health records.

Real-world evidence (RWE) generated from EHR-RWD hold unprecedented potential to bridge the unmet gaps that exist between controlled clinical trial studies and the complexities of healthcare delivery in the real-world [1,2,7,9]. While randomised control trial (RCT) studies remain the “gold standard” for assessing the efficacy of new interventions, the essence of RWE lies in its potential to reflect the diversity, heterogeneity, and nuances of patient populations and care settings, therefore enabling a more holistic understanding of health outcomes and interventions [7,10]. Data and evidence from real-world studies can support the life cycle of drug development, clinical and regulatory decision-making, and health technology assessment [3,8,11]. Moreover, RWE underpins the vision of the learning healthcare system (LHS), which is a paradigm built upon the cycle of continuous learning to achieve personalised medicine [12]. The transformative potential of RWE however, hinges on a pivotal caveat – high quality data.

Despite its potential, EHR-RWD, and by extension RWE, grapple with formidable barriers and chief among these is data quality (DQ) [1,2,9,10,13,14]. The need for high-quality data was exemplified during the COVID-19 pandemic, when EHR-RWD was critical for research and planning [15]. Regulatory authorities, such as the FDA (Food and Drug Administration), EMA (European Medicines Agency), NICE (National Institute for Health and Care Excellence), and MHRA (Medicines and Healthcare products Regulatory Agency), recommend the reporting of DQ metrics

and dimensions to provide additional context to real-world study outcomes, therefore serving as the foundation for trustworthy RWE [3,14,16–19]. These guidelines mostly promote ad-hoc DQ assessment and reporting, with the exception that EMA briefly notes the importance of assessing DQ as close as possible to the moment of data capture to help with collection errors [3,16–18]. While ad-hoc DQ measurement and reporting can increase transparency and awareness of limitations to real-world study outcomes, understanding the causes of poor data capture is needed for long term, sustainable improvement in DQ and, subsequently, the impact of RWE to support decision making in healthcare.

In this article, we aim to evaluate the robustness of studies seeking to improve the quality of structured EHR-derived data. Incorporating quality improvement best practices, we aim to assess how studies measure DQ, identify which interventions are implemented, and summarise the outcomes to understand which interventions are useful to improve DQ.

Data Quality Theory

The quality of data describes if data meets the expectations of a data consumer and therefore if it is fit for purpose [20]. This expected behaviour can be documented and understood using metadata – that is, additional data that provides meaning and context by describing how the data it is associated with should have been captured, defined, structured, and represented [20]. Metadata, in turn, can inform the design of quantifiable metrics that measure the compliance of data against a set of relevant business rules and constraints during DQ assessment [21]. DQ metrics, which can act as the method by which to measure the respective DQ dimensions, can serve crucial purposes in understanding areas needing improvement.

Scholars have proposed various multi-dimensional frameworks for comprehensive DQ assessment. While there may be disagreements on semantic choices and definitions across these frameworks, certain DQ concepts are consistently studied and well-represented in the literature, popular frameworks, and DQ profiling software [22]. Despite the challenge, some common concepts can be mapped across frameworks, as exemplified by the six core dimensions defined by the Data Management Association (DAMA): completeness, validity, consistency, uniqueness, timeliness, and accuracy (refer to Table 1) [20].

In an exploration of the DQ literature, we found 10 other reviews that summarise the most frequently represented DQ dimensions across DQ studies, software, and theoretical frameworks [19,21,23–30]. All 10 reviews demonstrated the concept of completeness to be well represented. Of these, 4 reviews agreed that data accuracy and consistency were popular DQ concepts [23,24,26,28]. Timeliness and validity were said to be popular by 3 different reviews, and uniqueness was only raised by Gordon et al. [29].

Completeness, also known as missingness, is often reported to be popular among DQ frameworks, tooling, and studies [22–27,29,30]. In general, it refers to the degree to which all required or expected data values or records are present [20]. The most common method to measure completeness consists of counting records with blank, unknown, empty, “NULL”, or “NaN” values, though variations may include a measurement of data availability [22,25,26]. Reviews by Weiskopf & Weng [22] and Syed et al. [28] also found variations of data completeness assessment involving triangulation of multiple sources to create a gold standard. This approach however risks assuming the accuracy of available data. Previous work demonstrate that missing data can lower statistical power of research outcomes and lead to biased assumptions with improper use of imputation methods [10,31,32].

Data accuracy is also well studied in DQ literature [22,25,28]. The accuracy dimension measures the extent to which data reflects the truth of events and conform to their actual value [20,22,25,28]. Other terms used to describe accuracy include error, correctness, integrity, trustworthiness, reliability, and validity [22,28]. The most common method to assess accuracy in healthcare involves the comparison of EHR data to a reference gold standard, which may include paper records, manual data reviews, triangulation of data from multiple sources, or interviews with patients [22]. Measurement of data accuracy can identify issues such as lack of specificity or precision [33]. Previous work found that code precision can be related to staff training and/or use of multiple EHR systems [33–36]. As Cook et al. [37] noted in a review of DQ issues affecting social determinants data, imprecise codified data may affect minority groups disproportionately, which in turn may affect secondary research outcomes.

Various terms and definitions for validity exist in the literature [22,28]. It generally describes the conformance of data to expected value ranges, patterns, formats, general medical knowledge, or data standards as set by local or external authorities [22,26,28,38]. Validity is also termed plausibility, conformance, and integrity, and can also be separated into internal and external validity or be incorporated with other data elements such as temporal validity [22,38]. Since EHR-derived data contains large volumes of categorical data, such as patient demographic, diagnostic, and treatment-related information, validity constraints are needed to identify areas needing standardisation [39,40]. Data standardisation has been shown to correlate positively to data sharing capabilities and emergency care [41]. However, improper design of standardised data entry user interfaces, such as the use of excessively long drop-down lists for diagnostic codes, can also increase cognitive demand, lower workflow efficiency, and correlate to clinician burnout [42].

Timeliness refers to several time-related characteristics of data and is therefore also termed currency, recency, or freshness [22,38]. For example, time-related data items can be used to measure how close to the event the data describes was the information recorded. Factors affecting timely capture of EHR-data include workflow inefficiencies, documentation burden, limited access to hardware, and interruptions [43–48]. Batch processing of data long after the event may indicate a lack of timeliness and affects other DQ dimensions such as completeness, accuracy, and validity [28,44,46,48,49].

Consistency, otherwise known as concordance, describes the agreement of similar data elements between multiple sources [20,22]. The existence of multiple data capture systems and RWD sources can give rise to inconsistent data for a given patient, and in the absence of a defined gold standard, the consistency dimension can identify potentially erroneous data [10,50]. Botsis et al. [36] identified multiple inconsistencies during a DQ analysis of a pancreatic cancer patient cohort stored in the Columbia University Medical Centre’s EHR data warehouse. These included pancreatitis recorded as chronic in pathology reports, but acute in clinical notes, and diabetes patients receiving both codes for type-1 and type-2 in the same EHR source [36]. Lucadou et al. [51] found similar discrepancies when comparing data items between different systems adding that inconsistencies may be caused by individual documentation habits. Measurement of data consistency highlights potential duplication and redundancy between different EHR sources and within the same EHR system and can help thus improve data capture and/or data engineering workflows.

The uniqueness dimension identifies where duplication of objects, events, or values are not expected [10,26]. Duplication of patient EHRs frequently occur when disparate data flows that contain overlapping objects are combined [10,38]. Like consistency, the uniqueness can identify and resolve redundant and inefficient workflows and processes [52]. This is particularly relevant given that 60% to 90% of clinicians routinely copy and paste data between systems [53]. The ‘copy and paste’

phenomenon is pervasive in healthcare, and is known to promote inconsistencies, propagate errors, and contribute to documentation burden and clinician fatigue [42,48,53]. As such, the uniqueness dimension is related to consistency and accuracy.

Table 1: Core dimensions of [a]DQ as defined by [b]DAMA [20]. This table outlines the six fundamental DQ dimensions defined by DAMA to enable standardisation and comparability of DQ concepts defined, assessed, and improved in the included DQ improvement literature. These dimensions are essential for assessing and improving the quality of [c]RWD.

DQ Dimension	Description
Completeness	The presence of the expected data.
Uniqueness	Uniqueness of records where duplication is not expected.
Timeliness	A measure of data freshness.
Consistency	A check of consistency between multiple sources of the same data elements.
Validity	The validity of data against data standards or plausible values, ranges, or patterns.
Accuracy	A check of consistency of source data against a reference “gold standard”.

[a]DQ: data quality
[b]DAMA: data management association
[c]RWD: real-world data

Data Quality Tooling

DQ measurement involves the process by which defective values are identified and labelled through the application of business rules or automated tooling [21]. The subsequent analysis of DQ results can then be aggregated, analysed, and summarised providing key insights for improvement. Tools to support these activities are widely reported and well-studied [21,27,29,30,54,55].

While the availability of DQ tools is abundant, the literature reveals a considerable gap in the effective support for DQ improvement efforts, particularly in the realm of health contexts. Evaluations of DQ profiling software by Ehrlinger & Wöß [30], Gordon et al. [29], and Ozonze et al. [21] highlight limitations in the range of DQ metrics offered for assessment, interoperability issues, and complex configuration requirements.

Root cause analysis, a pivotal aspect of DQ management, is also notably lacking in demonstrations within this landscape. Syed et al. [48] demonstrated the utility of the Odigos framework in qualitative root cause analysis, which classifies DQ issues that emanate from the material world, such as digital infrastructure or access to hardware, personal world, i.e. staff behaviours, and societal world, i.e. job roles and social norms. The legal and technical implications associated with data cleansing, as opposed to addressing the root causes of poor data, underscore potential risks to patient safety [56–58]. Consequently, the overarching trend in DQ tool development leans towards prioritising technical features, leaving a noticeable gap in the demonstration of their utility in the prevention and improvement of poor-quality data capture in real-world health settings.

Quality Improvement

Quality improvement describes the use of systematic continuous approaches to create positive changes in an area of need [59]. Various structured, iterative learning frameworks, such as Plan-Do-

Study-Act (PDSA), Total Data Quality Management (TDQM), DMAIC (Define-Measure-Analyse-Improve-Control), and the LHS, exist [12,60–63]. Lacking a universally agreed-upon model, each methodology focuses on enhancing different areas, ranging from service evaluation to treatment standards [59].

Developed from the earlier PDCA (Plan-Do-Check-Act) cycle by W. Edwards Deming, the PDSA cycle enhances the traditional model by prioritising the “study” stage – a deeper analysis rather than a simple check [64]. This adaptation roots the PDSA cycle firmly in the scientific method, encouraging a disciplined approach to testing and monitoring changes over time [64]. Its flexible and qualitative nature makes it particularly suitable for healthcare settings where adaptability to complex and variable processes is crucial.

Proposed by Richard Wang [61] in the late 1990s, TDQM adapts traditional Total Quality Management principles specifically to data management, highlighting the importance of data as a key asset or product. In healthcare, where decision-making increasingly relies on accurate and timely data, TDQM offers a robust framework to ensure the integrity and usability of data. This focus on DQ management is critical as healthcare systems integrate more digital processes and data-driven decision-making frameworks. TDQM adapts the PDSA planning stage to specifically target improvement of the quality of data [61].

A product of Motorola engineers in the 1980s, DMAIC provides a structured, data-driven quality improvement methodology [61]. Unlike the more qualitative PDSA, DMAIC emphasises quantifiable metrics and statistical analysis to identify and mitigate variations in processes. This makes DMAIC highly suitable for healthcare areas requiring high levels of measurement precision and control, such as clinical laboratories or any clinical process where outcomes need to meet high standards of care.

While all three methodologies share a structured, iterative approach and a reliance on empirical data to drive improvements, they cater to different needs within the healthcare sector. PDSA’s qualitative and flexible nature is best suited for areas requiring rapid change and adaptability. In contrast, DMAIC’s rigorous, statistical approach fits environments where precision and control are paramount. TDQM’s specific focus on data quality fills a critical niche in ensuring the reliability of healthcare data systems. What remains unknown is their implementations in the real-world. Taylor et al. [60] in a review of PDSA cycles aimed at improving treatment standards demonstrated that fewer than 20% of implementations comply with the core features including running multiple iterative learning cycles, the notion of small-scale change, and the use of quantitative data at monthly or more frequent intervals to inform progression of cycles. The stages of each methodology are listed and described in Table 2.

Table 2: Detailed comparison of three [a]QI frameworks for iterative learning. This table presents a comparison of the stages involved in three iterative learning frameworks present in health QI and other related literature. The comparison aims to help understand common themes in QI methodology, and how these can be applied to [b]DQ improvement.

Framework	Stage	Description
[c]PDSA [60]		
	Plan	Identify a change hypothesis and plan a small test.
	Do	Conduct a study plan with the collection of data.
	Study	Analyse and interpret the results.

	Act	Adapt the change based on feedback and plan the next iteration.
[d]TDQM [61]		
	Define	Define target data requirements and [c]DQ dimensions.
	Measure	Create metrics to evaluate these dimensions.
	Analyse	Investigate root causes for DQ issues.
	Improve	Identify key areas for improvement based on DQ root cause analysis.
[e]DMAIC [63]		
	Define	Define project scope and objectives.
	Measure	Identify and measure baseline service indicators.
	Act	Analyse baseline metrics and identify causes of errors.
	Improve	Implement changes to reduce or remove root causes of defects.
	Control	Put mechanisms in place to ensure sustained improvement.

[a]QI: quality improvement
[b]DQ: data quality
[c]PDSA: Plan-Do-Study-Act
[d]TDQM: total data quality management
[e]DMAIC: define measure act improve control

Study Objectives

The rapidly growing body of DQ publications and software tools indicate that this field has gained significant traction, and recent publications illustrates that there is no shortage of DQ concepts, frameworks, and tools [21,24,28–30,54,55,65,66]. While these surveys already provide comprehensive theoretical and functional evaluations on existing DQ concepts and tools for definition and measurement, this represents only the early stages of a bigger picture in DQ improvement and management. Our aim is to evaluate the robustness of studies seeking to use DQ measurement as part of DQ improvement initiatives, focusing on improving the quality of structured EHR-derived data.

Brouwer et al. [66], Wiebe et al. [67], and Lemma et al. [68] have previously published on DQ driven improvement in healthcare. These articles are compared in
. Wiebe et al. [67] included 24 studies aiming to improve EHR documents such as operative reports or discharge summaries. The authors reported that heterogeneity in tools or metrics used to measure the quality of unstructured clinical notes made it difficult to evaluate outcomes however, 8 included studies used an ad hoc questionnaire and 1 used the validated Physician Documentation Quality Instrument (PDQI-9) tool. Although unstructured notes in healthcare is a pervasive and ubiquitous source of important patient information, this scope limits the exploration of semi- or automated DQ assessment tools or methods [67].

Meanwhile, Brouwer et al. [66] evaluated studies published up to 2005 and limited to a general practice setting. With digital healthcare technology and culture evolving rapidly across healthcare settings and in the time since publication, a more recent and broader evaluation is needed. Lemma et al. [68] focused on low- and middle-income countries, where initiatives generally targeted broader

and less specific DQ improvement compared to high-income countries or technologically advanced institutions. The review expands on these works by evaluating contemporary DQ improvement studies targeting structured EHR-derived RWD agnostic of healthcare settings. Our evaluation is guided by quality improvement best practices to understand how studies measure and seek to improve DQ dimensions as defined by the well-recognised DAMA framework [20]. Specifically, we address the following three questions:

1. How do quality improvement studies define and measure the quality of data?
2. What interventions are implemented seeking to improve the quality of RWD?
3. What are the outcomes reported?

Table 3: Comparison of literature reviews evaluating [a]DQ improvement studies. This table aims to compare various literature reviews focused on DQ Improvement, identifying current knowledge gaps, and evaluating the existing body of research. The goal is to understand current progress and unmet needs.

Review Paper	Brouwer et al. [66]	Wiebe et al. [67]	Lemma et al. [68]
Time Period Covered	<2005	2004-2016	2008-2020
Number of Studies	12	24	20
Structured RWD[b]	x	-	x
Completeness	x	x	x
Accuracy	x	x	x
Timeliness	-	x	x
Consistency	-	-	x
Validity	-	-	-
Uniqueness	-	-	-
QI[c] Framework	-	-	-

[a]DQ: data quality
[b]RWD: real-world data
[c]QI: quality improvement
x: did evaluate
-: did not evaluate

Methods

Search Strategy and Information Sources

In this review, studies seeking to improve the quality of structured EHR data were examined using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [69]. The Population-Intervention-Comparison-Outcome-Context (PICOC) framework was used to identify relevant keywords and MeSH (Medical Subject Headings) terms [70,71]. These were combined using Boolean operators to generate strategic search queries that were implemented in the Ovid MEDLINE and Pubmed databases searching from 1945 to July 2023 (see Multimedia Appendix 1 for more information). Additional relevant papers were identified from other publications and manual searches through Google Scholar.

Literature Selection Process

The Ovid MEDLINE and Pubmed search results were downloaded as RIS and Pubmed files, respectively. These were then imported into Mendeley reference manager and Rayyan software for iterative analysis [72]. The Rayyan web-app was used to streamline the selection process. Articles were selected based on meeting all the following criteria: (1) describe a DQ assessment or measurement process, (2) target data from an EHR or EMR system, and (3) implemented an intervention seeking to improve DQ over time. The search strategies and article selection process were performed independently by two reviewers: AL and MA. Excluded articles were non-empirical studies, improvement studies focusing on quality of care or treatment standards instead of quality of data, studies targeting semi- or un-structured data or data not captured by a RWD source such as an EHR or EMR system, and studies without an intervention seeking to improve DQ. Error: Reference source not found summarises the key inclusion/exclusion criteria for paper selection. In total, 39 studies were included in the review, as presented by the PRISMA flow diagram in
below.

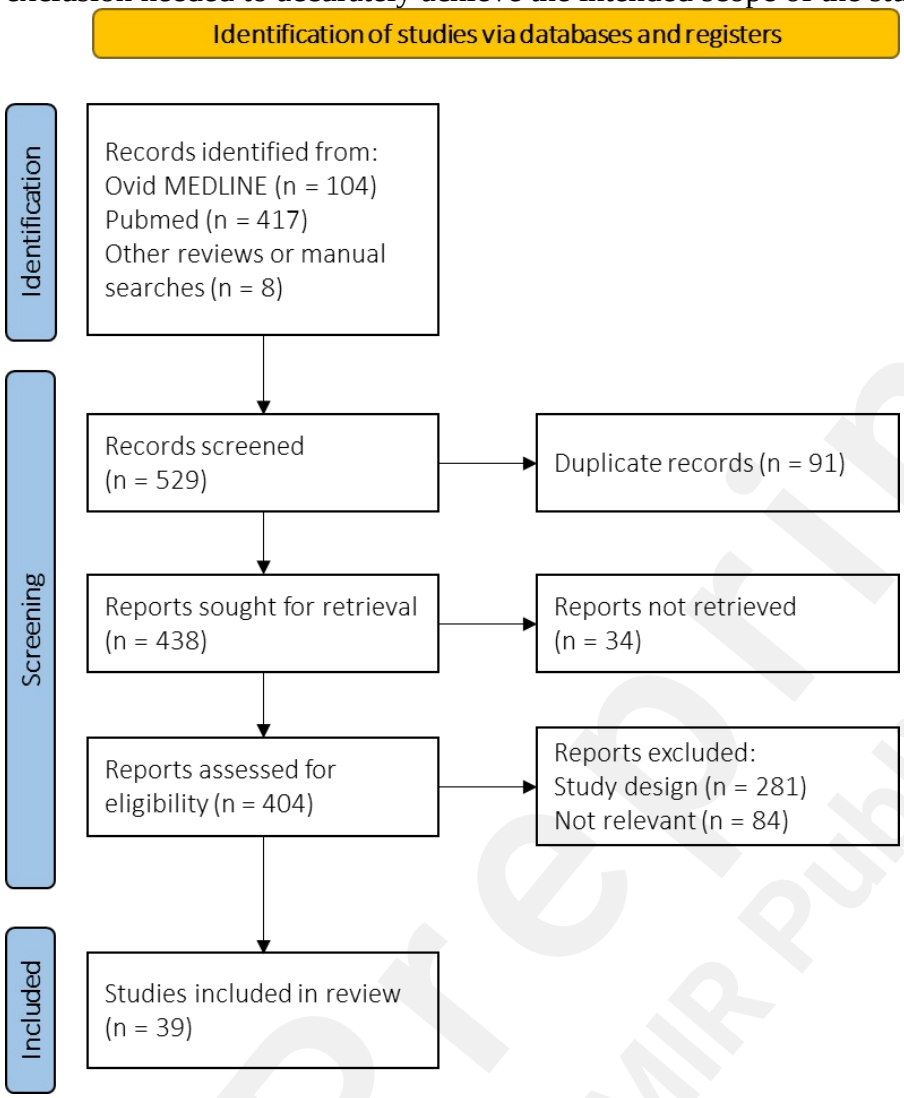
Table 4: Inclusion and exclusion criteria to evaluate the current landscape of [a]DQ assessment and improvement approaches. This table details the specific criteria used to include or exclude studies in the evaluation of current methods for assessing and improving the quality of structured health [b]RWD. These criteria help to systematically assess the landscape of DQ assessment and improvement strategies.

Criteria	Inclusion Criteria	Exclusion Criteria
Article Type	Empirical, original, or review articles where tools, frameworks, or interventions seek to measure and improve DQ.	Non-empirical studies, thesis papers, and non-peer-reviewed publications.
Language	Published in the English language.	Articles not published in English.
Access	Peer-reviewed articles	Papers that are not free to access
Primary Target for Quality Improvement	Studies that primarily aim to improve the quality of data.	Studies that primarily target improvement of treatment standards, standard of care, clinical workflows without a DQ focus.
Study Population	Studies targeting structured, tabular data.	Studies targeting semi- or un-structured data.
Data Source	Data from RWD sources such as EHR, EMR, PACS, or HIS-like systems.	Data generated by clinical trial studies.
DQ Assessment and Reporting	Studies that describe a DQ assessment, quantification, or measurement process and implemented an intervention seeking to improve DQ over time.	Studies that focus on DQ tool development without demonstration of measurement or improvement of DQ over time.
Location or Health Context	No criteria applied.	No criteria applied.
Time frame	Studies published from 1945 onwards.	Studies published before 1945.

[a]DQ: data quality

[b]RWD: real-world data
[c]EHR: electronic health record
[d]EMR: electronic medical record
[e]PACS: picture archiving and communication system.
[f]HIS: hospital information system

Figure 1: PRISMA flowchart for study selection to evaluate current literature targeting assessment, analysis, and improvement of the quality of structured [a]RWD in healthcare. This PRISMA flowchart illustrates the selection process for reports included in the review, detailing the steps of inclusion and exclusion needed to accurately achieve the intended scope of the study.



[a]RWD: real-world data

Data Extraction and Synthesis

Following paper selection, we defined a set of data elements essential for addressing the defined research questions. Subsequently, two authors (AL and MA) independently extracted and documented this information from each study. The collected data were then cross-checked for notable discrepancies, and any disparities were resolved through consensus. This is summarized below in Table 5 and Table 6 (more information available in Multimedia Appendix 2). Five key sections of information were extracted: (1) study characteristics (e.g. year of publication and health domain), (2) study plans outlined (e.g. descriptions of target data and relevant metadata to plan DQ assessment), (3) DQ assessment (e.g. methods and dimensions), (4) interventions (e.g. which interventions were implemented seeking to improve DQ?), and (5) outcomes (e.g. how are results reported?).

Results

Overview of Study Characteristics

We identified 39 studies describing DQ improvement initiatives in healthcare that targeted structured RWD sources [73–111]. These are listed in Table 5. Most were published from 2009 to 2022 (n = 37), with 2 published in 2002. Studies were affiliated with institutions in the US (n = 15), followed by Kenya (n = 4), Australia (n = 2), or one of 15 other countries. We classified the affiliated institutions into different levels of healthcare, including primary (n = 11), secondary (n = 10), tertiary (n = 15), and community (n = 3). The primary domains of healthcare were general practice (n = 4), HIV care (n = 4), intensive care (n = 4), tropical medicine (n = 4), oncology (n = 3), surgery (n = 3), or one of 12 other domains.

DQ improvement studies targeted RWD sources that were generated by varying numbers of institutions. Target data was generated by either at a single organisation (n = 9) or multiple different sites, which ranged from 2 to 10 (n = 11), 11 to 50 (n = 9), or over 51 (n = 9) different organisations. A variety of terminology were used to describe the source systems, including EMR (n = 12), national registries or databases (n = 8), EHR (n = 7), HIS (n = 7), clinical information networks (n = 6), PACS (n = 1), or claims data (n = 1).

Data Quality Assessment Methods

We found various approaches to DQ assessment. The duration of studies ranged from 1 month to 9 years as did the frequency of DQ assessment. Most measured and reported DQ before and after the intervention (n = 38) at varying intervals, including a single before-and-after comparison (n = 19), yearly (n = 4), quarterly (n = 2), monthly (n = 11), fortnightly (n = 1), weekly (n = 1), or specified data cycles (n = 1). DQ assessment was achieved using manual (n = 15), automated (n = 3), or semi-automatic (n = 13) methods, whereas some methods lacked sufficient description (n = 6) to be classified. Semi-automated methods for DQ assessment mostly involved the scheduling of ad-hoc, manually curated programmatic scripts using either R, SAS, or SQL programming languages [75,84,87,89,90,95–98,100,105–107]. Three studies applied automated methods which used the WHO DQ assessment tool and the Open Data Kit [81,90,93], but did not explicitly describe how this was implemented.

The approaches to defining and assessing DQ dimensions varied. We found that 6 studies explicitly referenced one of two DQ frameworks, including Weiskopf & Weng [22] (n = 3) and Kahn et al. [38] (n = 3) [78,79,81,87,89,93]. To enable comparison between studies, DQ metrics and dimensions reported were extracted and classified according to the DAMA DQ framework. Some DQ concepts lacked sufficient detail to allow classification (e.g. studies reporting “error rate”, “wrong data”, and “percentages of correctly coded” [92,95,111]). These were classified as “unclear” (n = 8). DQ improvement studies assessed the dimensions of completeness (n = 31), accuracy (n = 12), validity (n = 12), consistency (n = 11), timeliness (n = 6), and uniqueness (n = 4). The number of DQ dimensions targeted per study were one (n = 13), two (n = 14), three (n = 2), or four (n = 7).

Studies reported inconsistent terminology and definitions for DQ dimensions. For example, although the completeness dimension was generally assessed as the presence or absence of expected data, variations included the proportion of linkage of records between systems [108], use of a gold standard to identify missing patients [91], or overlapping completeness with other dimensions such as validity [75] or accuracy [107]. Validity, also termed conformance or plausibility, was targeted by 12 studies. Of these, 6 used data standards such as WHO ICD version 9 or 10 [76,84,88,89],

SNOMED [76,78,88], HL7 [76,87,88], or RxNorm [89]. Others assessed validity by defining business rules that incorporated expected values, formats, or ranges based on local or general medical knowledge [84,88,89,97]. Studies occasionally equated validity with accuracy or correctness [75,88,97].

Of 12 studies that targeted accuracy, 8 reported the development and/or use of a gold standard for reference. However, varying definitions for what studies deemed to be a "gold standard" were provided. This included paper charts [92,107], national data [90,109], a manually curated dataset [83,105], or manual validation by a trained, expert clinical coder [102]. For example, Rahbar's et al. [105] "gold standard" included 30 patient records that were manually abstracted by a team of experts that included a vascular neurologist clinician before comparing to national stroke registry records. In another study by Ahlbrandt et al. [95], the gold standard was described as "*the documented and encoded (using OPS-Classification, the German modification of ICPM) surgical procedure*", which could be interpreted either as a data standard or patient data in electronic or paper form.

Sometimes DQ dimensions were subsumed by another. For example, data could only be deemed accurate when it was both complete and correct across multiple data elements [104]. In the absence of what is deemed a "gold standard", data consistency was similarly assessed by comparison with paper records [76,82,93], multiple registers [85], or national data [81].

We found that studies' assessment of uniqueness and timeliness were generally consistent with the DAMA definitions. For uniqueness, studies assessed if records were unexpectedly duplicated, for example in primary keys [78] or patient names [94]. Likewise, timeliness was consistently assessed as the difference in time between point of data capture versus actual timing of events [73,74,76,80,81,103]. In contrast, some of the DQ concepts reported could not be classified according to the DAMA DQ framework, including simplicity, acceptability, flexibility, stability, usefulness [76], and conformance to a specified data model [89]. These were collectively classified as "other" (n = 3).

Interventions for Improvement

Studies varied in their approaches to plan and deliver DQ improvement interventions. Twenty studies reported using quantitative or qualitative data analysis before planning an intervention. Qualitative analysis involved assessment of clinical workflow inefficiencies through process mapping techniques or staff surveys [73,107]. In contrast, quantitative analysis involved an assessment of DQ with interpretations of possible root causes [94,95].

To understand the types of interventions studied, we identified five common themes, including DQ reporting and feedback (n = 24), IT-related or technical solutions (n = 21), training (n = 17), workflow (n = 5), or data cleaning (n = 3). All studies implemented at least one intervention with most implementing multiple interventions (n = 23). DQ reporting and feedback involved assessing DQ and sharing curated results to a specific stakeholder with the aim of encouraging improved data capture behaviour. These stakeholders included individual clinical staff or managers [86,97,104] or healthcare institutions as a whole [81,87,89,100,108,109].

Taggart et al. [97] implemented structured DQ reports combined with feedback sessions to improve the quality of EHR data in general practice settings. This approach leveraged regular assessments and direct feedback to practice managers to foster ongoing improvements in data recording practices, illustrating a practical application of DQ feedback mechanisms in a real-world healthcare setting [97]. Sinaiko et al. [86] in contrast studied peer-comparison feedback emails in a randomised

controlled study to assess its effectiveness on improving cancer stage data completeness, underscoring the importance of control groups in validating the impact of DQ interventions.

We found a range of subthemes under the IT-related or technical-based interventions. These improvements involved the introduction of a new or upgrade to a pre-existing electronic data capture system [76,80,92,98,101,102,107,110], the front-end user interfaces [73,77,95], or the back-end data flow processes [88,94]. Ahlbrandt et al. [95] introduced an intervention focusing on improving the graphical user interface (GUI) of anesthesia information management systems to enhance the validity of data captured. By shifting from dropdown lists to radio buttons, rearranging the GUI layout, and limiting user options to a set list, they aimed to reduce invalid data entry by making the interface more intuitive and compliant with data standards. This study exemplifies how interface design can directly influence data validity and highlights the impact of front-end modifications [95].

Technology based interventions often overlapped with training and workflow changes. Ewing et al. [102] implemented a browser-assisted clinical coding software along with training, which in turn improved efficiencies in clinical workflows. Other studies mainly targeted workflow inefficiencies [73,84,103,106] with Greiver et al. [106] introducing a data entry clerk whereas Moomba et al. [84] shifting data entry responsibility from data entry clerks to frontline clinical staff.

To plan, implement, and assess the impact of these DQ improvement interventions, we found that only 5 studies used a standardised quality improvement framework or iterative learning cycle, such as PDSA (n = 4) [73,80,83,100] or DMAIC [87]. All four PDSA studies completed multiple cycles, ranging from 3 to 8. One study reportedly conducted 421 PDSA cycles across 54 different sites [100]. Larrow et al. [73] applied the PDSA method to enhance the timeliness of discharge summaries at a paediatric hospital. The study team initiated their quality improvement project by identifying key barriers through qualitative analysis of staff surveys, leading to the strategic redesign of the EHR structured discharge summary template. Notable enhancements included embedded writing tips and standardised dropdown menus for common diagnoses.

Daniel et al. [87] applied the DMAIC methodology to define and assess DQ issues. These authors correlated specific DQ dimensions to possible technical issues, for example data lacking standardisation or valid entries may be caused by “*errors from data originators, ETL issues or limitations of the EHR data entry tool (inadequate value set constrains, lack of DQ checks)”* [87]. By measuring and analysing these problems in a structured methodology, the team identified key areas that required targeted interventions, such as use of data standards to enforce data validation rules to the data entry system [87].

Table 5: Standardised [a]DQ dimensions defined by [b]DAMA across the included studies. This table lists the six DQ dimensions defined by DAMA, as described, assessed, and/or improved in all included studied. It highlights the framework's role in standardising and comparing DQ concepts within the reviewed literature.

ID	Study	Completeness	Validity	Uniqueness	Consistency	Timeliness	Accuracy
1	De Lusignan (2002) [108]	x	x	-	x	x	-
2	Wallace (2002) [107]	x	-	-	-	-	x
3	Nassaralla (2009) [104]	x	-	-	-	-	x
4	Amoroso (2010) [111]	-	-	-	-	-	-
5	Greiver (2011) [106]	x	-	-	-	-	x
6	Ahlbrandt (2012) [95]	-	-	-	-	-	x
7	Mphatswe (2012) [109]	x	-	-	-	-	x
8	Rahbar (2013) [105]	x	x	-	x	-	x
9	Knight (2014) [100]	x	-	-	-	x	-
10	Siegel (2014) [99]	x	-	-	-	-	-
11	Benard (2015) [110]	x	-	-	-	-	-
12	Genet (2015) [103]	-	-	-	-	-	-
13	Haskew (2015) [98]	x	-	-	-	-	-
14	Smith (2015) [94]	x	-	x	-	-	-
15	Soto (2015) [101]	x	-	-	-	-	-
16	Taggart (2015) [97]	x	x	x	x	-	-
17	Ewing (2016) [102]	-	-	-	-	-	x
18	Ma (2016) [76]	x	x	-	x	x	-
19	Tuti (2016) [96]	x	-	-	-	-	-
20	Qin (2017) [92]	x	-	-	-	-	-
21	Edgerton (2018) [91]	x	-	-	-	-	-
22	Miyoshi (2018) [88]	x	x	x	-	-	-
23	Muthee (2018) [93]	x	-	-	x	-	-
24	Qualls (2018) [89]	x	x	-	-	-	-
25	Daniel (2019) [87]	x	x	x	x	-	-

26	Bhattacharaya (2020) [81]	x	-	-	x	x	x
27	Dean (2020) [74]	-	-	-	-	x	x
28	Koo (2020) [80]	-	-	-	-	-	-
29	Moomba (2020) [84]	x	x	-	-	-	-
30	Ng (2020) [77]	x	-	-	-	-	x
31	Njuguna (2020) [90]	-	x	-	-	-	-
32	Sinaiko (2020) [86]	x	-	-	-	-	-
33	Larrow (2021) [73]	x	-	-	-	-	x
34	Manesen (2021) [82]	x	-	-	x	-	-
35	Olagundoye (2021) [83]	-	-	-	-	-	x
36	Tizifa (2021) [85]	x	-	-	x	-	-
37	Kiogou (2022) [79]	x	x	-	x	x	-
38	Pfaff (2022) [78]	x	x	-	x	-	-
39	Tuti (2022) [75]	x	x	-	-	-	-
	Totals	31	12	4	11	6	12

[a]DQ: data quality

[b]DAMA: data management association

x: did assess and target the DQ dimension for improvement

-: did not assess or target the DQ dimension for improvement

Table 6: Themes of interventions for [a]DQ improvement in included studies. This table summarises the various intervention themes implemented to improve the quality of structured [b]RWD in various healthcare contexts. It provides insights into common strategies in DQ improvement.

ID	Study	[c]IT / Technical	Training	DQ Report & Feedback	Workflow	Cleaning
1	De Lusignan (2002) [108]	-	-	x	-	-
2	Wallace (2002) [107]	x	x	-	-	-
3	Nassaralla (2009) [104]	-	x	x	-	-
4	Amoroso (2010) [111]	x	x	x	-	-
5	Greiver (2011) [106]	-	x	-	x	-
6	Ahlbrandt (2012) [95]	x	x	-	-	-
7	Mphatswe (2012) [109]	-	x	x	-	-
8	Rahbar (2013) [105]	-	-	x	-	x
9	Knight (2014) [100]	x	x	x	-	-
10	Siegel (2014) [99]	-	-	x	-	-
11	Benard (2015) [110]	x	-	-	-	-
12	Genet (2015) [103]	-	-	-	x	-
13	Haskew (2015) [98]	x	-	x	-	-
14	Smith (2015) [94]	x	-	-	-	-
15	Soto (2015) [101]	x	x	x	-	-
16	Taggart (2015) [97]	-	-	x	-	-
17	Ewing (2016) [102]	x	x	-	-	-
18	Ma (2016) [76]	x	-	-	-	-
19	Tuti (2016) [96]	-	x	x	-	-
20	Qin (2017) [92]	x	-	-	x	-
21	Edgerton (2018) [91]	-	-	-	-	-
22	Miyoshi (2018) [88]	x	-	-	-	-
23	Muthee (2018) [93]	x	-	x	-	-
24	Qualls (2018) [89]	-	-	x	-	-
25	Daniel (2019) [87]	x	x	x	-	-
26	Bhattacharaya (2020) [81]	-	x	x	-	-
27	Dean (2020) [74]	x	-	-	-	-
28	Koo (2020) [80]	x	x	x	-	-
29	Moomba (2020) [84]	-	-	-	x	-
30	Ng (2020) [77]	x	x	x	-	x
31	Njuguna (2020) [90]	-	-	x	-	-
32	Sinaiko (2020) [86]	-	-	x	-	-
33	Larrow (2021) [73]	x	x	-	x	-
34	Manesen (2021) [82]	x	-	x	-	-
35	Olagundoye (2021) [83]	x	x	x	-	-
36	Tizifa (2021) [85]	x	x	x	-	-

37	Kiogou (2022) [79]	-	-	-	-	x
38	Pfaff (2022) [78]	-	-	x	-	-
39	Tuti (2022) [75]	-	-	x	-	-
	Totals	21	17	24	5	3

[a]DQ: data quality
[b]RWD: real world data
[c]IT: information technology
x: did implement
-: did not implement

Reported Outcomes

To understand and compare the outcomes of DQ improvement initiatives, we identified whether studies reported DQ changes that were better, worse, or no change over time. Most studies reported improvements in DQ over time (n = 36). This excludes three studies due to results being reported as preliminary [87], potential improvements as opposed to actual [91], or without sufficient detail [78]. Of 36 studies showing improvement, 9 also report decreases in DQ [95,99,105,108,110], of which another 4 also report no changes [81,84,97]. These changes were reported at varying levels of granularity. While most studies reported DQ metrics for specific data items, such as validity of surgical procedure codes [95], others aggregated multiple metrics or dimensions into higher level entities, such as, "*92% reduction in error rate*" [111] and mean monthly accuracy for PEWS scores [74].

We also assessed whether statistical tests were used to demonstrate significance of effect and whether studies compared intervention groups with a control group. When determining significance of treatment effect, 22 studies used at least one statistical test or method [73–75,77,80–82,86,92,93,95,97–99,102–106,108,109]. These ranged from chi-squared (n = 7) and statistical process control charts (n = 4) to multivariable linear (n = 1) and logistic regression (n = 1). Of these, only 1 study compared the intervention group to a concurrent control group who reported improvement in completeness of cancer stage data [86].

Discussion

Principal Findings

In this paper, we conducted a systematic literature review to understand current practices in DQ improvement of structured RWD in a healthcare context. We found substantial heterogeneity in the approaches to definition, assessment, and interventions across the reviewed literature. The range of definitions for DQ concepts, quality improvement methodologies, and reported outcomes have made synthesis and comparison of the results challenging. Below we explore these three points in greater depth.

Data Quality

A key issue in the exploration of DQ is the lack of consensus on theoretical definitions for DQ assessment. Despite the existence of several DQ frameworks, there are no agreed recommendations or guidelines on which frameworks should be employed, or on how dimensions should be defined, measured, or used to understand real-world issues in data capture, processing, and utility for high quality RWE generation. This has been demonstrated in a wealth of previous reviews on DQ theory [22,23,25,26,38,48], but to a lesser extent in a quality improvement context. In our review, we found that while some studies did in fact reference theoretical frameworks by Weiskopf and Weng [22] and

Kahn et al. [38], these account for fewer than 15% of all included studies. This indicates a severe lack of uptake of standardised DQ theory in the wider literature and explains the substantial variation and lack of consensus. In turn, the lack of agreement and consistency makes it difficult to harness the true purpose of DQ assessment, which pertains to its ability to identify issues in real-world processes, behaviours, and resources. While some studies demonstrate qualitative correlations between DQ issues and underlying real-world problems [48], we found only a small minority of studies implemented a quantitative approach to make a similar connection.

Since DQ is a complex, multidimensional construct, each dimension serves to identify context-specific issues in the real-world needing remediation. We found that some authors made this correlation either directly or indirectly. For example: data validity is affected by a lack of standardisation of front-end user interfaces on electronic data capture forms [95]; timeliness of data indicates possible workflow inefficiencies that delays the point of data capture [73,103]; duplication highlights redundant data sources [94]; inaccurate data underpins lack of training on medical coding standards [102], and inconsistencies between data sources indicates possible capture of inaccurate data [82,85]. This raises two important points: the need to assess DQ beyond completeness or missingness, and the importance of standardised frameworks. Without these, crucial error-prone processes in complex clinical pathways may go undiagnosed and continue to generate poor quality data. This is particularly important given the growing demand for and expectations of real-world healthcare data, the hype in artificial intelligence (AI), and the growing awareness that maintaining patient records is the leading cause of clinician burnout [42,112].

There was limited reporting of the tools or software used for DQ assessment included in this review. Only two DQ assessment tools were reported: the WHO DQ assessment toolkit [81,90,93] and the ‘Open Data Kit’ [85]. Neither of these are explained in sufficient depth to discern how they work or their applicability to other environments. Further investigation into the referenced material also lacked sufficient information. Other tools reported were scheduled programmatic scripts using R, SQL, or SAS software for DQ assessment. Some of these methods are considered ‘automated’ solutions for DQ assessment. This indicates a significant gap between the vast range of DQ software available and the practical implementation of these tools for DQ assessment, causal analysis, and improvement. DQ software must be capable of profiling large volumes of structured data, provide both automated and user specified DQ assessment methods, and facilitate meaningful analysis of possible root causes of poor DQ [21,30]. The limited adoption of existing DQ software might suggest a deficiency in technical proficiency, inadequate documentation clarifying its utility or use cases, or a lack of awareness regarding its availability or relevance.

Quality Improvement Cycles

Another characteristic of the studies included in this review was the limited use of quality improvement frameworks. In only 5 studies [73,80,83,87,100] was a quality improvement methodology referenced to plan and implement DQ improvement interventions. This is surprising given the potential benefits these frameworks offer, particularly in fostering systematic, structured, and dynamic approaches to improvement in complex environments.

Quality improvement frameworks like PDSA, DMAIC, and TDQM, if implemented robustly, can significantly improve comparability and knowledge sharing between studies, institutions, and organisational teams. This is important given that Siegel et al. [99] observed varying improvements across different organisations, stating that systematic and organised quality improvement efforts are needed.

However, the strengths of these frameworks extend beyond iterative learning; they also encourage a

deep dive into DQ analysis, helping to unravel the complex relationships between various real-world factors and the root causes of poor DQ. In this way, interventions can be designed in collaboration with the affected stakeholders, i.e. frontline clinical staff, to maximise the opportunity for DQ improvement. Knight et al. [100] particularly emphasised this point stating that the quality improvement model, i.e. PDSA, "*used in this project facilitated the identification and correction of difficulties with the technology of the innovation"*. Quality improvement frameworks, such as PDSA, TDQM, or DMAIC, can also be adapted to improve the quality of real-world healthcare data incorporating DQ driven quantitative analysis alongside real-world issues that can be identified using the Odigos framework [48].

Despite their strengths, the application of these frameworks is not without challenges. One significant constraint is the need for substantial upfront planning and stakeholder engagement, which can be resource intensive. Furthermore, these frameworks require a culture of continuous improvement and openness to change and adoption of data governance practices, which may not be present in all healthcare settings. This can limit their applicability and effectiveness. Additionally, the lack of consistent application and reporting on the use of these frameworks can make it difficult to evaluate their true effectiveness.

Outcomes

We sought to investigate current approaches to DQ assessment and improvement to synthesise and summarise the lessons learnt from these endeavours. In general, studies reported positive changes in DQ through the implementation of multiple interventions. Lemma et al. [68] associated the benefits to DQ when interventions such as training, technical innovation, and DQ feedback were combined. The same authors reported that studies that focused only on single interventions did not generate equally positive DQ changes. In contrast, we found that 17 studies focused on a single intervention showing mostly positive results. For example, Sinaiko et al. [86] demonstrated the positive impact of peer-comparison emails to completion of cancer stage data when compared to a control group. Likewise, studies that demonstrated a combination of improvements, reductions, and no changes in DQ often implemented multiple interventions.

Studies highlighted the importance of close collaboration with clinical users' needs when implementing digital technologies [100,107,111] and the importance of detailed, personalised feedback on data capture performance provided to data capture end-users [86,106,108]. In contrast, Taggart et al. [97] found that peer-comparison and feedback sessions did not result in better DQ and suggested the need for randomised controlled studies. This raises an important question regarding the need for control groups in DQ improvement studies. In our review, we found only one study that compared an intervention group with a concurrent control group yet most report successful improvement in DQ.

While some studies highlight the need for controlled environments, we observed that interventions were chosen for reasons other than baseline quantitative or qualitative analysis or analysis of underlying causes. In other words, only a few studies planned an intervention based on a data-driven approach. When comparing outcomes, most studies compared average baselines pre- and post-intervention, where only one study compared the intervention group to a concurrent non-intervention control group. A simultaneous control group for comparison can facilitate analysis of the cause of effect of treatment interventions along with the implementation of multiple interventions, which occurred in most studies. This combined with the lack of data-driven approaches when planning or designing interventions indicates a significant gap in robust, standardised DQ improvement methodologies. Therefore, the reported outcomes should be considered with caution.

The methodological rigor in DQ improvement studies often suffers from a lack of randomised controls and consistent statistical methodologies. As discussed above, only one reviewed study included a randomised control group to demonstrate significant improvement [86], and just 22 out of 39 studies applied diverse statistical tests like chi-squared and logistic regression. This inconsistency in applying robust analytical techniques can introduce biases, misattributing improvements to interventions rather than actual effects. Moreover, the absence of uniform experimental designs across various healthcare settings undermines the robustness and generalisability of findings.

Incorporating structured methodologies such as PDSA or DMAIC could significantly enhance the methodological rigor of these studies. These frameworks support systematic implementations and evaluations, facilitating the use of control groups and statistical analysis to reliably isolate intervention effects. By adopting such standardised approaches, future research could more effectively ensure the credibility and applicability of findings, fostering the development of evidence-based interventions suitable for diverse healthcare environments.

Future Recommendations

This review highlights the need for standardised and systematic approaches to DQ assessment, analysis, and improvement. This can be addressed in future studies by following quality improvement methodologies, such as the PDSA [113], TDQM [61], or DMAIC [63] iterative learning cycles, and DQ frameworks, such as DAMA [20], Weiskopf et al. [22], or Kahn et al. [38]. Furthermore, understanding the root causes of poor DQ is essential for planning the most appropriate intervention which should aim to address issues as close as the point of data capture as possible.

The need for standardised DQ assessment is evident. Future research and development should focus on the development and demonstration of DQ tools that are not only grounded in theoretical frameworks like those offered by DAMA [20], Weiskopf et al. [22], or Kahn et al. [38] but are also highly accessible and user-friendly. DQ tools should come with comprehensive documentation and practical examples that enable users in making informed decisions about their applicability and relevance in specific healthcare settings.

As discussed above, DQ tooling currently lack in useability and usefulness [21,29,30]. Ease of use can be overcome by introducing “plug-and-play” functionality that is combined with useful customisable features. This duality can allow users to quickly test and assess the tool’s immediate value and adjust and extend its functionality to fit more complex, specific needs over time. By extension, DQ tools should produce results that are useful for meaningful, in-depth analysis and monitoring of DQ errors.

Current best practices in root cause analysis of poor real-world healthcare data are unknown but could be facilitated using a framework such as the Odigos framework [48]. Furthermore, understanding causes of poor data can facilitate the design and selection of more relevant interventions needed – an aspect of DQ management that was demonstrated by few articles in this review. Future studies may also wish to compare intervention groups to concurrent control groups and explore one intervention at a time instead of multiple. This may help to control for external factors and increase understanding of barriers to high quality data capture.

Limitations

This review aimed to summarise the lessons learnt from DQ improvement studies. Since an abundance of literature already highlights the substantial variation in terminology for DQ concepts,

we utilised the DAMA DQ framework to standardise the heterogeneity in DQ terms and definitions. In doing so, some DQ concepts could not be classified, therefore potentially affecting the frequency counts of DQ dimensions assessed. Another limitation is that we were unable to perform a comprehensive meta-analysis of the methodological constrains and the effect measures of the reported outcomes. We believe the significant scope of this work warrants future work. This was due to considerable variation in the methods for assessment, analysis, and reporting of DQ metrics and changes over time. This posed significant challenges when attempting to objectively elucidate the effect of treatment interventions.

Moreover, while this review captures a rise in DQ improvement studies, with 31 out of 39 studies published in the last decade, it also includes 8 studies that were published in or before 2013, potentially missing recent advancements in digital health technologies. Additionally, the discovery of 6 additional studies from manual searches indicates the likely exclusion of other relevant work. This is particularly the case due to the lack of consistency in DQ terminology and definitions, which made it difficult to capture all possible variations of DQ terms in the search strategy. Despite these challenges, our review included twice as many studies compared to other related reviews, indicating a thorough coverage within the constraints identified. We believe the significant scope of this work warrants future updates to include emerging trends and methodologies in DQ improvement.

Conclusion

The reviewed studies demonstrate that approaches to DQ improvement vary in their methodologies, definitions, and reporting of DQ dimensions. In general, studies implemented multiple interventions and reported positive changes in the quality of structured real-world healthcare data. In addition to ‘going paperless’ initiatives, studies demonstrated the benefits of engagement with frontline clinical end users, provision of personalised DQ feedback, streamlining clinical workflows, and raising awareness of DQ and data standards aimed at improving DQ in healthcare settings. Despite this, heterogeneity is a major limitation among DQ literature in general, and we recommend that studies refer to standardised frameworks, such as PDSA cycles for quality improvement and the DAMA DQ framework for assessing DQ dimensions. This would lead to greater consistency and comparison in reported outcomes.

Funding

This work was supported by Cancer Research UK RadNet Manchester [C1994/A28701]. Dr Price was supported by the NIHR Manchester Biomedical Research Centre (NIHR203308).

Acknowledgements

The authors would like to thank Dr Catherine O’Hara, Head of Analytics from the Clinical Outcomes and Data Unit for her support and guidance in this project, as well as Daniel Livesey, Librarian at The Christie NHS Foundation Trust, for his invaluable training and assistance on defining and implementing the literature search strategy.

Data Availability

The datasets containing all results extracted from the studies used to write this paper are provided as Multimedia Appendices in the journal submission.

Conflicts of Interest

None declared.

Abbreviations

AI: artificial intelligence
DAMA: Data Management Association
DMAIC: Define Measure Analyse Improve Control
DQ: data quality
EHR: electronic health records
EMA European Medicines Agency
EMR: electronic medical records
FDA: Food and Drug Administration
HIS: hospital information system
HL7: health language 7
ICD: International Classification for Disease
IT: information technology
LHS: learning healthcare system
MeSH: Medical Subject Headings
MHRA: Medicines and Healthcare Products Regulatory Agency
NICE: National Institute for Health and Care Excellence
PACS: picture archiving and communication systems
PDQI-9: Physician Documentation Quality Instrument
PDSA: Plan Do Study Act
PEWS: paediatric early warning scores
PICOC: Patient Intervention Comparison Outcome Context
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RWD: real-world data
RWE: real-world evidence
SNOMED: systematized nomenclature of medicine
SQL: structured query language
TDQM: Total Data Quality Management
WHO: World Health Organisation

References

1. Knevel R, Liao KP. From real-world electronic health record data to real-world results using artificial intelligence. *Ann Rheum Dis*. 2022;0:1-6. doi:10.1136/annrheumdis-2022-222626
2. Liu F, Demosthenes P. Real-world data: a brief review of the methods, applications, challenges and opportunities. *BMC Med Res Methodol*. 2022;22(1). doi:10.1186/s12874-022-01768-6
3. Real-World Evidence | FDA. Published 2023. Accessed June 21, 2023. https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence
4. Atasoy H, Greenwood BN, Mccullough JS. The Digitization of Patient Care: A Review of the Effects of Electronic Health Records on Health Care Quality and Utilization. *Annual Review of Public Health Annu Rev Public Health*. 2019;13(1):487-500. doi:10.1146/annurev-publhealth
5. Benson T. *Information in Practice Why General Practitioners Use Computers and Hospital Doctors Do Not-Part 1: Incentives*.; 2001. doi:10.1136/bmj.325.7372.1086
6. Anshari M. Redefining Electronic Health Records (EHR) and Electronic Medical Records (EMR) to Promote Patient Empowerment. *IJID (International Journal on Informatics for Development)*. 2019;8(1):35. doi:10.14421/ijid.2019.08106
7. Blonde L, Khunti K, Harris SB, Meizinger C, Skolnik NS. Interpretation and Impact of Real-World Clinical Data for the Practicing Clinician. *Adv Ther*. Published online October 24, 2018. doi:10.6084/m9.figshare.7156850
8. Guinn D, Wilhelm EE, Lieberman G, Khozin S. Assessing function of electronic health records for real-world data generation. *BMJ Evid Based Med*. 2019;24(3):95-98. doi:10.1136/bmjebm-2018-111111
9. Kim HS, Lee S, Kim JH. Real-world evidence versus randomized controlled trial: Clinical research based on electronic medical records. *J Korean Med Sci*. 2018;33(34). doi:10.3346/jkms.2018.33.e213
10. Dang A. Real-World Evidence: A Primer. *Pharmaceut Med*. 2023;37(1):25-36. doi:10.1007/s40290-022-00456-6
11. *NICE Real-World Evidence Framework Corporate Document*.; 2022. www.nice.org.uk/corporate/ecd9
12. Price G, Mackay R, Aznar M, et al. Learning healthcare systems and rapid learning in radiation oncology: Where are we and where are we going? *Radiotherapy and Oncology*. 2021;164:183-195. doi:10.1016/j.radonc.2021.09.030
13. Zhang J, Symons J, Agapow P, et al. Best practices in the real-world data life cycle. *PLOS Digital Health*. 2022;1(1):e0000003. doi:10.1371/journal.pdig.0000003
14. Blacketer C, Defalco FJ, Ryan PB, Rijnbeek PR. Increasing trust in real-world evidence through evaluation of observational data quality. *Journal of the American Medical Informatics Association*. 2021;28(10):2251-2257. doi:10.1093/jamia/ocab132
15. Thygesen JH, Tomlinson C, Hollings S, et al. COVID-19 trajectories among 57 million adults in England: a cohort study using electronic health records. *Lancet Digit Health*. 2022;4(7):e542-e557. doi:10.1016/S2589-7500(22)00091-7
16. NICE. Overview | NICE real-world evidence framework | Guidance | NICE. Published June 23, 2023. Accessed June 21, 2023. https://www.nice.org.uk/corporate/ecd9/chapter/overview
17. EMA. *Data Quality Framework for EU Medicines Regulation*.; 2022. www.ema.europa.eu/contact
18. GOV.UK. MHRA guidance on the use of real-world data in clinical studies to support regulatory decisions. Published 2021. Accessed October 9, 2023. https://www.gov.uk/government/publications/mhra-guidance-on-the-use-of-real-world-data-in-clinical-studies-to-support-regulatory-decisions/mhra-guidance-on-the-use-of-real-world-data-in-clinical-studies-to-support-regulatory-decisions#contents
19. Castellanos EH, Bradley MC, Forshee RA, Rivera DR. Raising the Bar for Real-World Data in

Oncology : Approaches to Quality Across Multiple Dimensions. *JCO Clin Cancer Inform*. 2024;8. doi:10.1200/CCI.23.00046
20. Earley S, Henderson D, Sebastian-Coleman L, eds. *DAMA*. 2nd ed. Technics Publications; 2017.
21. Ozonze O, Scott PJ, Hopgood AA. Automating Electronic Health Record Data Quality Assessment. *J Med Syst*. 2023;47(1). doi:10.1007/s10916-022-01892-2
22. Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: Enabling reuse for clinical research. *Journal of the American Medical Informatics Association*. 2013;20(1):144-151. doi:10.1136/amiajnl-2011-000681
23. Liaw ST, Rahimi A, Ray P, et al. Towards an ontology for data quality in integrated chronic disease management: A realist review of the literature. *Int J Med Inform*. 2013;82(1):10-24. doi:10.1016/j.ijmedinf.2012.10.001
24. Cichy C, Rass S. An overview of data quality frameworks. *IEEE Access*. 2019;7:24634-24648. doi:10.1109/ACCESS.2019.2899751
25. Chen H, Hailey D, Wang N, Yu P. A review of data quality assessment methods for public health information systems. *Int J Environ Res Public Health*. 2014;11(5):5170-5207. doi:10.3390/ijerph110505170
26. Bian J, Lyu T, Loiacono A, et al. Assessing the practice of data quality evaluation in a national clinical data research network through a systematic scoping review in the era of real-world data. *Journal of the American Medical Informatics Association*. 2020;27(12):1999-2010. doi:10.1093/jamia/ocaa245
27. Mariño J, Kasbohm E, Struckmann S, Kapsner LA, Schmidt CO. R Packages for Data Quality Assessments and Data Monitoring: A Software Scoping Review with Recommendations for Future Developments. *Applied Sciences (Switzerland)*. 2022;12(9). doi:10.3390/app12094238
28. Syed R, Eden R, Makasi T, et al. Digital Health Data Quality Issues: Systematic Review. *J Med Internet Res*. 2023;25. doi:10.2196/42615
29. Gordon B, Fennessy C, Varma S, et al. Evaluation of freely available data profiling tools for health data research application: a functional evaluation review. *BMJ Open*. 2022;12(5):e054186. doi:10.1136/bmjopen-2021-054186
30. Ehrlinger L, Wöß W. A Survey of Data Quality Measurement and Monitoring Tools. *Front Big Data*. 2022;5. doi:10.3389/fdata.2022.850611
31. Parry MG, Sujenthiran A, Cowling TE, et al. Imputation of missing prostate cancer stage in English cancer registry data based on clinical assumptions. *Cancer Epidemiol*. 2019;58:44-51. doi:10.1016/j.canep.2018.11.003
32. Westerberg M, Beckmann K, Gedeborg R, et al. Choice of imputation method for missing metastatic status affected estimates of metastatic prostate cancer incidence. *J Clin Epidemiol*. 2023;155:22-30. doi:10.1016/j.jclinepi.2022.12.008
33. Daniels B, Havard A, Myton R, Lee C, Chidwick K. Evaluating the accuracy of data extracted from electronic health records into MedicineInsight, a national Australian general practice database. *Int J Popul Data Sci*. 2022;7(1). doi:10.23889/IJPDS.V7I1.1713
34. Palestine AG, Merrill PT, Saleem SM, Jabs DA, Thorne JE. Assessing the Precision of ICD-10 Codes for Uveitis in 2 Electronic Health Record Systems. *JAMA Ophthalmol*. 2018;136(10):1186-1190. doi:10.1001/jamaophthalmol.2018.3001
35. Horsky J, Drucker EA, Ramelson HZ, &women's Hospital B. *Accuracy and Completeness of Clinical Coding Using ICD-10 for Ambulatory Visits*.; 2018. Accessed October 18, 2023. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5977598/
36. Botsis T, Hartvigsen G, Chen F, Weng C. Secondary Use of EHR: Data Quality Issues and Informatics Opportunities. Published online 2010. Accessed October 18, 2023. https://pubmed.ncbi.nlm.nih.gov/21347133/
37. Cook LA, Sachs J, Weiskopf NG. The quality of social determinants data in the electronic health record: a systematic review. *Journal of the American Medical Informatics Association*.

2022;29(1):187-196. doi:10.1093/jamia/ocab199
38. Kahn MG, Callahan TJ, Barnard J, et al. A Harmonized Data Quality Assessment Terminology and Framework for the Secondary Use of Electronic Health Record Data. *eGEMs (Generating Evidence & Methods to improve patient outcomes)*. 2016;4(1):18. doi:10.13063/2327-9214.1244
39. Bright CJ, Lawton S, Benson S, et al. Data Resource Profile: The Systemic Anti-Cancer Therapy (SACT) dataset. *Int J Epidemiol*. 2020;49(1):15-15L. doi:10.1093/ije/dyz137
40. Sarwar T, Seifollahi S, Chan J, et al. The Secondary Use of Electronic Health Records for Data Mining: Data Characteristics and Challenges. *ACM Comput Surv*. 2023;55(2). doi:10.1145/3490234
41. Zhang J, Ashrafian H, Delaney B, Darzi A. Impact of primary to secondary care data sharing on care quality in NHS England hospitals. *NPJ Digit Med*. 2023;6(1). doi:10.1038/s41746-023-00891-y
42. Diaz-Garelli F, Strowd R, Ahmed T, et al. *What Oncologists Want: Identifying Challenges and Preferences on Diagnosis Data Entry to Reduce EHR-Induced Burden and Improve Clinical Data Quality*. Vol 5.; 2021. https://doi.org/10.
43. Ullah E, Albrett J, Khan O, et al. Workload involved in vital signs-based monitoring & responding to deteriorating patients: A single site experience from a regional New Zealand hospital. *Heliyon*. 2022;8(10). doi:10.1016/j.heliyon.2022.e10955
44. Pérez-Martí M, Casadó-Marín L, Guillén-Villar A. Electronic Records With Tablets at the Point of Care in an Internal Medicine Unit: Before-After Time Motion Study. *JMIR Hum Factors*. 2022;9(1). doi:10.2196/30512
45. Byrd JB, Vigen R, Plomondon ME, et al. Data quality of an electronic health record tool to support VA cardiac catheterization laboratory quality improvement: The VA Clinical Assessment, Reporting, and Tracking System for Cath Labs (CART) program. *Am Heart J*. 2013;165(3):434-440. doi:10.1016/j.ahj.2012.12.009
46. Skyttberg N, Vicente J, Chen R, Blomqvist H, Koch S. How to improve vital sign data quality for use in clinical decision support systems? A qualitative study in nine Swedish emergency departments. *BMC Med Inform Decis Mak*. 2016;16(1). doi:10.1186/s12911-016-0305-4
47. Dall'Ora C, Griffiths P, Hope J, et al. How long do nursing staff take to measure and record patients' vital signs observations in hospital? A time-and-motion study. *Int J Nurs Stud*. 2021;118. doi:10.1016/j.ijnurstu.2021.103921
48. Eden R, Syed R, Makasi T, et al. Revealing the Root Causes of Digital Health Data Quality Issues: A Qualitative Investigation of the Odigos framework. In: Bui E, ed. 56th Hawaii International Conference on System Sciences (HICCS), Hyatt Regency Maui; 2023:3035-3044.
49. Sockolow PS, Bowles KH, Adelsberger MC, Chittams JL, Liao C. Impact of homecare electronic health record on timeliness of clinical documentation, reimbursement, and patient outcomes. *Appl Clin Inform*. 2014;5(2):445-462. doi:10.4338/ACI-2013-12-RA-0106
50. Lee D, Jiang X, Yu H. Harmonized representation learning on dynamic EHR graphs. *J Biomed Inform*. 2020;106. doi:10.1016/j.jbi.2020.103426
51. Von Lucadou M, Ganslandt T, Prokosch HU, Toddenroth D. Feasibility analysis of conducting observational studies with the electronic health record. *BMC Med Inform Decis Mak*. 2019;19(1). doi:10.1186/s12911-019-0939-0
52. Zayas-Cabán T, Haque SN, Kemper N. Identifying Opportunities for Workflow Automation in Health Care: Lessons Learned from Other Industries. *Appl Clin Inform*. 2021;12(3):686-697. doi:10.1055/s-0041-1731744
53. Tsou AY, Lehmann CU, Michel J, Solomon R, Possanza L, Gandhi T. Safe practices for copy and paste in the EHR: Systematic review, recommendations, and novel model for Health IT collaboration. *Appl Clin Inform*. 2017;8(1):12-34. doi:10.4338/ACI-2016-09-R-0150
54. Quindroit P, Fruchart M, Degoul S, et al. Definition of a practical taxonomy for referencing data quality problems in healthcare databases. *Methods Inf Med*. Published online 2022. doi:10.1055/a-1976-2371
55. Liaw ST, Guo JGN, Ansari S, et al. Quality assessment of real-world data repositories across the data

life cycle: A literature review. *Journal of the American Medical Informatics Association*. 2021;28(7):1591-1599. doi:10.1093/jamia/ocaa340
56. Shi X, Prins C, Van Pottelbergh G, Mamouris P, Vaes B, De Moor B. An automated data cleaning method for Electronic Health Records by incorporating clinical knowledge. *BMC Med Inform Decis Mak*. 2021;21(1). doi:10.1186/s12911-021-01630-7
57. Phan HTT, Borca F, Cable D, Batchelor J, Davies JH, Ennis S. Automated data cleaning of paediatric anthropometric data from longitudinal electronic health records: protocol and application to a large patient cohort. *Sci Rep*. 2020;10(1). doi:10.1038/s41598-020-66925-7
58. Dziadkowiec O, Callahan T, Ozkaynak M, Reeder B, Welton J. Using a Data Quality Framework to Clean Data Extracted from the Electronic Health Record: A Case Study. *eGEMs (Generating Evidence & Methods to improve patient outcomes)*. 2016;4(1):11. doi:10.13063/2327-9214.1201
59. Backhouse A, Ogunlayi F. Quality improvement into practice. *The BMJ*. 2020;368. doi:10.1136/bmj.m865
60. Taylor MJ, Mcnicholas C, Nicolay C, Darzi A, Bell D, Reed JE. Systematic review of the application of the plan-do-study-act method to improve quality in healthcare. *BMJ Qual Saf*. 2014;(23):290-298. doi:10.1136/bmjqs
61. Wang RY. A Product Perspective on Total Data Quality Management. *Commun ACM*. 1998;41(2).
62. D'Andreamatteo A, Ianni L, Lega F, Sargiacomo M. Lean in healthcare: A comprehensive review. *Health Policy (New York)*. 2015;119(9):1197-1209. doi:10.1016/j.healthpol.2015.02.002
63. Sokovic M, Pavletic D, Kern Pipan K. *Quality Improvement Methodologies-PDCA Cycle, RADAR Matrix, DMAIC and DFSS Industrial Management and Organisation*. Vol 43.; 2010. www.journalamme.org
64. Moen R, Norman C. *Evolution of the PDCA Cycle*.; 2009. https://www.anforq.org/activities/congresses/index.html
65. Burns L, Roux N Le, Kalesnik-Orszulak R, et al. Real-World Evidence for Regulatory Decision-Making: Guidance From Around the World. *Clin Ther*. 2022;44(3):420-437. doi:10.1016/j.clinthera.2022.01.012
66. Brouwer HJ, Bindels PJE, Van Weert HC. Data quality improvement in general practice. *Fam Pract*. 2006;23(5):529-536. doi:10.1093/fampra/cml040
67. Wiebe N, Otero Varela L, Niven DJ, Ronksley PE, Iragorri N, Quan H. Evaluation of interventions to improve inpatient hospital documentation within electronic health records: A systematic review. *Journal of the American Medical Informatics Association*. 2019;26(11):1389-1400. doi:10.1093/jamia/ocz081
68. Lemma S, Janson A, Persson LÅ, Wickremasinhe D, Källestål C. Improving quality and use of routine health information system data in low- And middle-income countries: A scoping review. *PLoS One*. 2020;15(10 October). doi:10.1371/journal.pone.0239683
69. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *The BMJ*. 2021;372. doi:10.1136/bmj.n71
70. Carrera-Rivera A, Larrinaga F, Lasa G. Context-awareness for the design of Smart-product service systems: Literature review. *Comput Ind*. 2022;142. doi:10.1016/j.compind.2022.103730
71. Yang H, Lee HJ. Research trend visualization by MeSH terms from PubMed. *Int J Environ Res Public Health*. 2018;15(6). doi:10.3390/ijerph15061113
72. Rayyan - AI Powered Tool for Systematic Literature Reviews. Accessed June 30, 2023. https://www.rayyan.ai/
73. Larrow A, Chong A, Robison T, et al. A Quality Improvement Initiative to Improve Discharge Timeliness and Documentation. In: *Pediatric Quality and Safety*. Vol 6. Wolters Kluwer Health; 2021:E440. doi:10.1097/pq9.0000000000000440
74. Dean NP, Cheng JJ, Crumbley I, Duval J, Maldonado E, Ghebremariam E. Improving Accuracy and Timeliness of Nursing Documentation of Pediatric Early Warning Scores. *Pediatr Qual Saf*. 2020;5(2):E278. doi:10.1097/pq9.0000000000000278

75. Tuti T, Aluvaala J, Chelangat D, 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. Asweto CO, ed. *PLOS Global Public Health*. 2022;2(10):e0000673. doi:10.1371/journal.pgph.0000673
76. Ma S, Lawpoolsri S, Soonthornworasiri N, et al. Effectiveness of implementation of electronic malaria information system as the national malaria surveillance system in Thailand. *JMIR Public Health Surveill*. 2016;2(1). doi:10.2196/publichealth.5347
77. Ng C, Payne AS, Patel AK, Thomas-Mohtat R, Maxwell A, Abo A. Improving Point-of-Care Ultrasound Documentation and Billing Accuracy in a Pediatric Emergency Department. *Pediatr Qual Saf*. 2020;5(4):E315. doi:10.1097/pq9.0000000000000315
78. Pfaff ER, Girvin AT, Gabriel DL, et al. Synergies between centralized and federated approaches to data quality: A report from the national COVID cohort collaborative. *Journal of the American Medical Informatics Association*. 2022;29(4):609-618. doi:10.1093/jamia/ocab217
79. Kiogou SD, Chi CL, Zhang R, Ma S, Adam TJ. *Clinical Data Cohort Quality Improvement: The Case of the Medication Data in The University of Minnesota's Clinical Data Repository*.
80. Koo JK, Moyer L, Castello MA, Arain Y. Improving Accuracy of Handoff by Implementing an Electronic Health Record-generated Tool: An Improvement Project in an Academic Neonatal Intensive Care Unit. *Pediatr Qual Saf*. 2020;5(4):E329. doi:10.1097/pq9.0000000000000329
81. Bhattacharya AA, Allen E, Umar N, et al. Improving the quality of routine maternal and newborn data captured in primary health facilities in Gombe State, Northeastern Nigeria: A before-and-after study. *BMJ Open*. 2020;10(12). doi:10.1136/bmjopen-2020-038174
82. Manesen R, Mekler KA, Molobi TR, et al. Data quality assessment of a South African electronic registry for drug-resistant TB, 2015–2016. *Public Health Action*. 2021;11(1):33-39. doi:10.5588/pha.20.0031
83. Olagundoye O, Van Boven K, Daramola O, Njoku K, Omosun A. Improving the accuracy of ICD-10 coding of morbidity/mortality data through the introduction of an electronic diagnostic terminology tool at the general hospitals in Lagos, Nigeria. *BMJ Open Qual*. 2021;10(1). doi:10.1136/bmjoq-2020-000938
84. Moomba K, Williams A, Savory T, et al. Effects of real-time electronic data entry on HIV programme data quality in Lusaka, Zambia. *Public Health Action*. 2020;10(1):47-52. doi:10.5588/pha.19.0068
85. Tizifa TA, Nkhono W, Mtengula S, van Vugt M, Munn Z, Kabaghe AN. Leveraging phone-based mobile technology to improve data quality at health facilities in rural Malawi: a best practice project. *Malar J*. 2021;20(1). doi:10.1186/s12936-021-03742-x
86. Sinaiko AD, Barnett ML, Gaye M, Soriano M, Mulvey T, Hochberg E. Association of Peer Comparison Emails With Electronic Health Record Documentation of Cancer Stage by Oncologists. *JAMA Netw Open*. 2020;3(10):e2015935. doi:10.1001/jamanetworkopen.2020.15935
87. Daniel C, Serre P, Orlova N, Bréant S, Paris N, Griffon N. Initializing a hospital-wide data quality program. The AP-HP experience. *Comput Methods Programs Biomed*. 2019;181. doi:10.1016/j.cmpb.2018.10.016
88. Miyoshi NSB, De Azevedo-Marques JM, Alves D, De Azevedo-Marques PM. An eHealth platform for the support of a brazilian regional network of mental health care (eHealth-Interop): Development of an interoperability platform for mental care integration. *JMIR Ment Health*. 2018;5(4). doi:10.2196/10129
89. Qualls LG, Phillips TA, Hammill BG, et al. Evaluating Foundational Data Quality in the National Patient-Centered Clinical Research Network (PCORnet®). *eGEMs (Generating Evidence & Methods to improve patient outcomes)*. 2018;6(1):3. doi:10.5334/egems.199
90. Njuguna C, Vandi M, Mugagga M, et al. Institutionalized data quality assessments: A critical pathway to improving the accuracy of integrated disease surveillance data in Sierra Leone. *BMC Health Serv Res*. 2020;20(1). doi:10.1186/s12913-020-05591-x

91. Edgerton JR, Herbert MA, Hamman BL, Ring WS. Can use of an administrative database improve accuracy of hospital-reported readmission rates? *Journal of Thoracic and Cardiovascular Surgery*. 2018;155(5):2043-2047. doi:10.1016/j.jtcvs.2017.11.071
92. Qin Y, Zhou R, Wu Q, et al. The effect of nursing participation in the design of a critical care information system: a case study in a Chinese hospital. *BMC Med Inform Decis Mak*. 2017;17(1):165. doi:10.1186/s12911-017-0569-3
93. Muthee V, Bochner AF, Osterman A, et al. The impact of routine data quality assessments on electronic medical record data quality in Kenya. *PLoS One*. 2018;13(4). doi:10.1371/journal.pone.0195362
94. Smith BL, Ramadan M, Corley B, Hablas A, Seifeldein IA, Soliman AS. Measuring the effect of improvement in methodological techniques on data collection in the Gharbiah population-based cancer registry in Egypt: Implications for other Low- and Middle-Income Countries. *Cancer Epidemiol*. 2015;39(6):1010-1014. doi:10.1016/j.canep.2015.11.001
95. Ahlbrandt J, Henrich M, Hartmann BA, et al. Small cause - Big effect: Improvement in interface design results in improved data quality - A multicenter crossover study. In: *Studies in Health Technology and Informatics*. Vol 180. IOS Press; 2012:393-397. doi:10.3233/978-1-61499-101-4-393
96. Tuti T, Bitok M, Malla L, 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*. 2016;1(1). doi:10.1136/bmjgh-2016-000028
97. Taggart J, Liaw ST, Yu H. Structured data quality reports to improve EHR data quality. *Int J Med Inform*. 2015;84(12):1094-1098. doi:10.1016/j.ijmedinf.2015.09.008
98. Haskew J, Rø G, Saito K, et al. Implementation of a cloud-based electronic medical record for maternal and child health in rural Kenya. *Int J Med Inform*. 2015;84(5):349-354. doi:10.1016/j.ijmedinf.2015.01.005
99. Siegel EM, Jacobsen PB, Lee JH, et al. Florida initiative for quality cancer care: Improvements on colorectal cancer quality of care indicators during a 3-year interval. *J Am Coll Surg*. 2014;218(1). doi:10.1016/j.jamcollsurg.2013.09.008
100. Knight AW, Szucs C, Dhillon M, Lembke T, Mitchell C. The eCollaborative: Using a quality improvement collaborative to implement the national ehealth record system in Australian primary care practices. *International Journal for Quality in Health Care*. 2014;26(4):411-417. doi:10.1093/intqhc/mzu059
101. Soto M, Capurro D, Catalán S. Evaluating the data completeness in the Electronic Health Record after the Implementation of an Outpatient Electronic Health Record. In: *Studies in Health Technology and Informatics*. Vol 216. IOS Press; 2015:885. doi:10.3233/978-1-61499-564-7-885
102. Ewing M, Funk GA, Warren AM, et al. Improving National Trauma Data Bank® coding data reliability for traumatic injury using a prospective systems approach. *Health Informatics J*. 2016;22(4):1076-1082. doi:10.1177/1460458215610896
103. Genet IC, Firestone KS, Volsko TA. Neonatal respiratory therapist-led rounds can improve staff satisfaction and timeliness of respiratory interventions. *Respir Care*. 2015;60(3):321-327. doi:10.4187/respcare.03489
104. Nassaralla CL, Naessens JM, Hunt VL, et al. Medication reconciliation in ambulatory care: Attempts at improvement. *Qual Saf Health Care*. 2009;18(5):402-407. doi:10.1136/qshc.2007.024513
105. Rahbar MH, Gonzales NR, Ardjomand-Hessabi M, et al. The University of Texas Houston Stroke Registry (UTHSR): Implementation of enhanced data quality assurance procedures improves data quality. *BMC Neurol*. 2013;13. doi:10.1186/1471-2377-13-61
106. Greiver CCFP Director M, Toronto Research Network N, Professor A, et al. *Associate Family Physician-in-Chief, Chair of the Research Ethics Board, North York General Hospital*. Vol 19. British Computer Society; 2011.
107. Wallace CJ, Stansfield' D, Gibb Ellis" KA, Clemmerl TP. *Implementation of an Electronic Logbook*

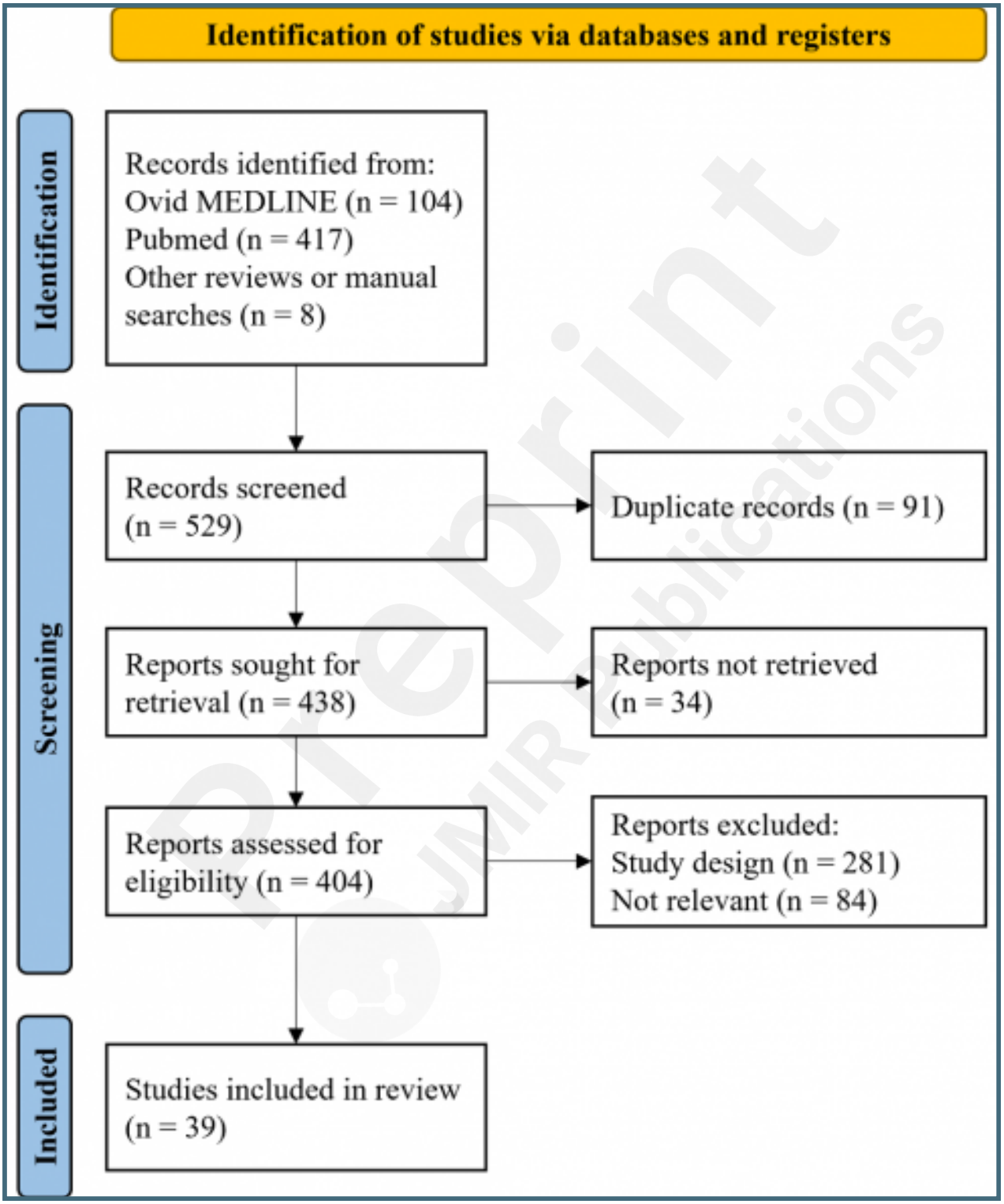
for Intensive Care Units.
108. De Lusignan S, Stephens PN, Adal N, Majeed A. Does feedback improve the quality of computerized medical records in primary care? In: *Journal of the American Medical Informatics Association*. Vol 9. ; 2002:395-401. doi:10.1197/jamia.M1023
109. Mphatswe W, Mate KS, Bennett B, et al. Amélioration des informations sur la santé publique: Une intervention sur la qualité des données dans le Kwazulu-Natal (afrique du sud). *Bull World Health Organ*. 2012;90(3):176-182. doi:10.2471/BLT.11.092759
110. Benard M, Bouaud J, Marsault C, Boudghene F, Carette MF, Séroussi B. Diagnostic Imaging Requisition Quality When Using an Electronic Medical Record: A Before-After Study. In: *Studies in Health Technology and Informatics*. Vol 210. IOS Press; 2015:364-368. doi:10.3233/978-1-61499-512-8-364
111. Amoroso CL, Akimana B, Wise B, Fraser HSF. Using electronic medical records for HIV care in rural Rwanda. In: *Studies in Health Technology and Informatics*. Vol 160. IOS Press; 2010:337-341. doi:10.3233/978-1-60750-588-4-337
112. Johnson KB, Neuss MJ, Detmer DE. Electronic health records and clinician burnout: A story of three eras. *Journal of the American Medical Informatics Association*. 2021;28(5):967-973. doi:10.1093/jamia/ocaa274
113. Christoff P. Running PDSA cycles. *Curr Probl Pediatr Adolesc Health Care*. 2018;48(8):198-201. doi:10.1016/j.cppeds.2018.08.006

Last page intentionally left blank

Supplementary Files

Figures

PRISMA flowchart for study selection to evaluate current literature targeting assessment, analysis, and improvement of the quality of structured aRWD in healthcare. This PRISMA flowchart illustrates the selection process for reports included in the review, detailing the steps of inclusion and exclusion needed to accurately achieve the intended scope of the study.



Multimedia Appendixes

Search Strategy & Results.
URL: http://asset.jmir.pub/assets/2a0179625dd345ba188ef26b80cf5bf2.docx

Data Items Extracted.
URL: http://asset.jmir.pub/assets/99f341e37d0d544671a7459883ec8301.docx

All Studies Results.
URL: http://asset.jmir.pub/assets/1a741947ceb7b5c3951b6d9aec7e3066.docx

PRISMA Checklist.
URL: http://asset.jmir.pub/assets/1120ef7afbcb87f09aa2d340343d7440.docx