Review

Electronic health record data quality assessment and tools: a systematic review

Abigail E. Lewis[1,2], **Nicole Weiskopf** [iD][3], **Zachary B. Abrams** [iD][2], **Randi Foraker** [iD][2], **Albert M. Lai** [iD][2], **Philip R.O. Payne** [iD][2], and **Aditi Gupta**[2,*]

[1]Division of Computational and Data Sciences, Washington University in St. Louis, St. Louis, Missouri, USA
[2]Institute for Informatics, Data Science and Biostatistics, Washington University in St. Louis, St. Louis, Missouri, USA
[3]Department of Medical Informatics and Clinical Epidemiology, Oregon Health & Science University, Portland, Oregon, USA

*Corresponding Author: Aditi Gupta, PhD, Institute for Informatics, Data Science and Biostatistics, Washington University in St. Louis, 660 S. Euclid Ave., Campus Box 8132, Saint Louis, MO 63110, USA; agupta24@wustl.edu

ABSTRACT

Objective: We extended a 2013 literature review on electronic health record (EHR) data quality assessment approaches and tools to determine recent improvements or changes in EHR data quality assessment methodologies.

Materials and Methods: We completed a systematic review of PubMed articles from 2013 to April 2023 that discussed the quality assessment of EHR data. We screened and reviewed papers for the dimensions and methods defined in the original 2013 manuscript. We categorized papers as data quality outcomes of interest, tools, or opinion pieces. We abstracted and defined additional themes and methods though an iterative review process.

Results: We included 103 papers in the review, of which 73 were data quality outcomes of interest papers, 22 were tools, and 8 were opinion pieces. The most common dimension of data quality assessed was completeness, followed by correctness, concordance, plausibility, and currency. We abstracted conformance and bias as 2 additional dimensions of data quality and structural agreement as an additional methodology.

Discussion: There has been an increase in EHR data quality assessment publications since the original 2013 review. Consistent dimensions of EHR data quality continue to be assessed across applications. Despite consistent patterns of assessment, there still does not exist a standard approach for assessing EHR data quality.

Conclusion: Guidelines are needed for EHR data quality assessment to improve the efficiency, transparency, comparability, and interoperability of data quality assessment. These guidelines must be both scalable and flexible. Automation could be helpful in generalizing this process.

Key words: clinical research informatics, data quality, electronic health records

BACKGROUND

The usage of electronic health record (EHR) derived data in biomedical research has increased in recent years, and this trend is expected to continue as such technologies improve.[1] The multitude of data available in EHRs make them well-suited for high-dimensional analyses, including phenotyping as well as machine learning and artificial intelligence approaches.[2] Additionally, EHR data offer potential cost and time savings as an alternative to the primary collection of medical data for research purposes.[3] The coronavirus disease 2019 (COVID-19) pandemic, for example, has highlighted the importance of using EHR data to uncover and monitor patterns in disease spread and severity.[4,5] In addition to the potential benefits of using EHR data for research, however, there are also challenges, including EHR data quality concerns, timely access, patient protections and confidentiality, and the ability to generalize results based on EHR data.[3,4,6]

Best practices for assessing EHR data quality, despite their clear importance, remain an open question. There does not currently exist a standard approach to assessing EHR data quality and quality assessments are often ad hoc for a specific

project. Some work has been done previously to consolidate data quality assessment (DQA) approaches. Prior to the broad adoption and usage of EHRs, Wang and Strong presented a conceptual framework of data quality.[7] Although not specific to EHRs or medicine at all, their framework is still applicable in an EHR setting. Categories of data quality included in this framework are intrinsic data quality (data are objective and accurate), contextual data quality (quality is based on the context in which data were collected; data are relevant, timely, and complete), representational data quality (data are represented consistently and interpretable), and accessible data quality (data are accessible and securely managed).[7] The data consumers who helped develop this model indicated accessibility and context as some of the most important facets of data quality. Working under this mindset, DQAs as they relate to EHR data should be defined on their own accord.

A 2013 review established 5 themes of EHR data quality and 7 methods by which to assess these dimensions.[8] Dimensions included completeness (the presence of data in the EHR); correctness (the truthfulness of data in the EHR);

concordance (the agreement between elements within the EHR and between the EHR and other data sources); plausibility (the extent to which EHR data make sense in a larger medical context); and currency (the accuracy of the EHR data for the time at which it was recorded and how up to date the data are). At the time, concordance and plausibility seemed likely to be proxies for accuracy or correctness. Often, plausibility was defined by the correctness of a value in the EHR or the believability of a distribution of values in light of other knowledge. In some cases, plausibility implied a value is possible in the given setting without asserting the correctness of the value which may be prone to some sort of recording error. Assessment methods included a gold standard comparison between the EHR data and another data set that is considered to be true; data element agreement: agreement between elements within the EHR; element presence of necessary data fields and observations in the EHR; data source agreement: agreement between the EHR and another data source not necessarily considered to be a gold standard; distribution comparison of EHR data distributions to clinical data source distributions; validity checking of EHR data; and log review: an examination of data entry practices. Although the definition of certain methods and dimensions are similar, dimensions represented an element of data quality while methods represented the process used to assess for dimensions. Additionally, individual methods could be used to assess multiple dimensions of data quality. In this case, the dimension being assessed was determined by the rationale for the chosen method.

Additional data quality frameworks have been proposed beyond the 2013 review. Kahn et al highlight conformance, completeness, and plausibility in their proposed DQA framework, which has been utilized by groups like the National Patient-Centered Clinical Research Network (PCORnet) and the All of Us research program to streamline DQA.[9–11] Wang et al[12] propose a rule-based system for assessing data quality. The Observational Health Data Sciences and Informatics (OHDSI) program has developed the Automated Characterization of Health Information at Largescale Longitudinal Evidence Systems (ACHILLES), a framework and tool, for assessing data conforming to the observational medical outcomes partnership (OMOP) common data model (CDM).[13] Although not considered a universal tool, this tool can be openly accessed and used to assess the quality of data conforming to the OMOP structure. The utility of ACHILLES has been demonstrated in multiple settings for creating comparable DQAs.[14] The National COVID Cohort Collaborative (N3C) uses a version of the OHDSI DQA approach to assess data quality after transforming data to the OMOP CDM.[15] The dimensions and methods from the 2013 review and more recent attempts to standardize DQA, although consistently utilized across health-related research applications to describe data quality, still vary substantially in the ways they are recorded and discussed, indicating a lack of community agreement and adoption of such DQA frameworks and methods.

OBJECTIVE

To address this gap in knowledge, we aim to extend the 2013 literature review[8] to understand how EHR DQA has changed in recent years. We propose 2 main extensions of the 2013 review: first, to expand the literature review on EHR DQA to present day to determine how DQA practices have changed since 2013, and second, to include a broader range of literature including developed DQA tools and opinion pieces in addition to papers that focus on a data quality measure as an outcome of interest.

MATERIALS AND METHODS

As we were extending the prior literature review, we closely followed the methods outlined by Weiskopf.[8] Similarly, we aimed to identify articles which discussed the quality of EHR data specifically. In order to do this, we included the same terms as Weiskopf in our PubMed Query of titles and abstracts completed in September 2021:

> ((data accuracy[MeSH Major Topic]) OR ("data quality") OR ("data reliability") OR ("data validity") OR ("data error") OR ("data errors")) AND ("electronic health records"[MeSH Major Topic] OR "medical records systems, computerized"[MeSH Major Topic] OR "electronic medical record"[All Fields] OR "computerized medical record"[All Fields] OR "EMH"[Title/Abstract] OR "EHR"[Title/Abstract]) AND English[lang]

This PubMed query resulted in a total of 593 articles. To select articles for review, we developed the inclusion and exclusion criteria in Table 1. We developed inclusion and exclusion criteria such that they identified original work related to DQA of EHR data. These articles were then sorted in descending order by their number of citations per year since being published and reviewed in order until selecting 90 articles, a similar number to the 2013 review.[8] However, one of our goals was to capture both highly relevant articles and emergent literature. We first selected at least 10 articles from both 2020 and 2021 based on a descending sorting of number of citations per year before sorting and searching through the remainder of the query. We screened the abstracts of 253 papers and based inclusion on the criteria from Table 1. From here, we read 122 papers in full to determine inclusion which resulted in inclusion of a total of 90 papers in the review (Figure 1). We repeated the PubMed query in April 2023 to extend our search window through March 2023. An

Table 1. Inclusion and exclusion criteria

Inclusion criteria

Criterion	Reason for inclusion
Original work with data quality measure as outcome of interest	To examine specific approaches to EHR DQA
Original development of DQA tools	To gauge tools that are being used in healthcare fields to assess quality of EHR data
Opinion paper	To understand how EHR DQA is understood by experts in the field

Exclusion criteria

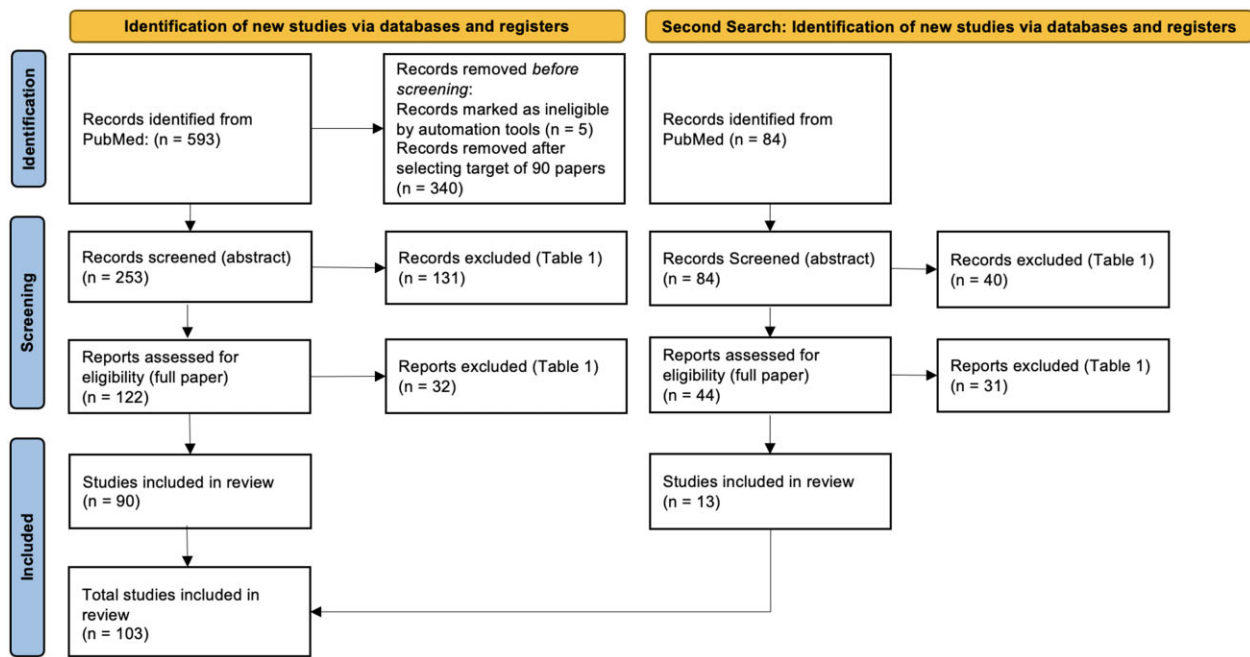| Criterion	Reason for exclusion
Review papers	Review papers do not employ or develop DQA tools and do not constitute original findings
Original work with a nondata quality measure as the outcome of interest	These papers do not specifically address issues of data quality

Figure 1. Prisma diagram.[16]

additional 84 articles were identified, of which 13 were included in the final collection based on meeting inclusion criteria (Table 1, Figure 1). Two authors (AEL, AG) developed inclusion and exclusion criteria to reduce bias during the screening process. A doctoral student (AEL) completed the screening and review analysis. A second author validated the selection of included papers and review analysis (AG).

For each included paper, we determined a paper category, the data quality dimensions assessed, and the methods used to assess the data quality dimensions. Paper categories included data quality outcomes of interest, tools, and opinion pieces. Data quality outcome of interest papers included original research using DQA methods. DQA tool papers included specific methodologies for assessing data quality or a set of definitions for understanding DQA, instructions for how to use them, and demonstration of the tool on one or more example datasets. Tool papers differed from data quality outcome of interest papers in that they are designed to be used on a general data set rather than a specific data set of interest. Opinion papers represented an amalgamation of DQA suggestions from experts in the field and differ from tools in that they do not include a tangible output and have not necessarily been tested.

We abstracted the presence of data quality dimensions as defined by Weiskopf[8] from each paper along with the methods used to assess the dimensions based on the definitions provided in the introduction. In addition, we collected the type of data being analyzed, vocabulary used to describe data quality dimensions, and specific evaluation methods within the larger methodological groups. We then used an iterative process to abstract and define additional dimensions and methods in all pieces and themes occurring in tools and opinion pieces as these themes may differ from those defined in 2013. In order to abstract new dimensions and methods during the first round of review, we recorded dimensions of data quality and methods that did not fit into one of Weiskopf's definitions. Commonly occurring topics were considered to

be new dimensions, methods, or themes. Themes encompassed all concepts not considered to be a method or dimension of data quality. We created the minimum number of mutually exclusive themes which included all commonly occurring concepts. We then reviewed all papers a second time for data collection on the newly defined dimensions, methods, and themes. A list of papers and the collected data can be found in the Supplementary Materials.

RESULTS

Of the 103 papers included in the review, 73 were data quality outcome of interest papers, 22 were tools, and 8 were opinion pieces (Table 2). Ninety-nine papers discussed structured data, 25 papers discussed unstructured data, and 21 discussed both. Table 3 and Figure 2 show the types of methods used to assess each dimension.

Similar to 2013, the most commonly assessed dimension of data quality was completeness which was explored in 76 (74%) papers (Table 2). In the majority of cases, element presence was used to assess completeness (Table 3). When completeness was assessed by comparison to another data set within or external to the EHR, gold standard or otherwise, comparison data sets included within EHR agreement,[17–23] an alternative data source,[18,24,25] billing data,[22] or physician agreement.[19,22,26,27] Common terms used to describe completeness included missingness, presence, availability, breadth, and accuracy.

Again, the second most commonly addressed dimension of data quality was correctness which was assessed in 53 (51%) papers (Table 2). Data element agreement was the method most often used to assess correctness and was followed closely by a gold standard comparison and data source agreement (Table 3). Comparison data sets included other EHR data from the same system,[18,28–37] manual review by a physician,[19,22,33,38–50] billing data,[22,51,52] unstructured data, or an external data source.[18,25,53–62] Many terms were used to

Table 2. Dimensions of data quality

Dimension	All papers	DQA outcome of interest	Tools	Opinion pieces
Total	103	73	22	8
Structured data	99 (96%)	70 (96%)	21 (95%)	8 (100%)
Unstructured data	25 (24%)	20 (27%)	2 (9%)	3 (38%)
Completeness	76 (74%)	50 (68%)	19 (86%)	7 (88%)
Correctness	53 (51%)	35 (48%)	10 (45%)	8 (100%)
Concordance	46 (45%)	36 (50%)	7 (32%)	3 (38%)
Plausibility	29 (28%)	15 (21%)	11 (50%)	3 (38%)
Currency	35 (34%)	19 (26%)	8 (36%)	8 (100%)
Conformance	18 (17%)	7 (10%)	8 (36%)	3 (38%)
Bias	11 (11%)	11 (15%)	0 (0%)	0 (0%)

describe correctness including accuracy, validity, specificity, sensitivity, positive predictive value, and error.

Forty-six (45%) papers assessed concordance most often using data element agreement or data source agreement (Tables 2 and 3). Similar to correctness, comparison data sets included other EHR data from the same system,[17,18,25,30,31,33,46,50,53,63–69] unstructured data,[34,55,70–73] manual review by a physician,[38,41,42,46,48,74–76] billing data,[46,68,69,77–80] or an external data source.[18,25,53,54,56,57,60,80–87] Common terms used to describe concordance were consistency, agreement, sensitivity, discrepancy, and correlation.

In contrast to 2013, currency was the fourth most commonly assessed dimension of data quality and was considered in 35 (34%) papers (Table 2). Currency was most often assessed using log review while data element agreement and distribution comparison were also utilized (Table 3). Common terms used to describe currency included timeliness, frequency, and accuracy. Finally, plausibility was assessed in 29 (28%) papers and was most often assessed by distribution comparison. Common terms used to describe plausibility included validity, truthfulness, extreme values, duplication, and believability.

In addition to the 5 dimensions of data quality identified by Weiskopf, we identified conformance and bias as further dimensions of data quality and structural agreement as a method by which to conduct DQA. Eighteen (17%) papers assessed conformance, or compliance with a predefined representational structure, almost exclusively using structural agreement (Tables 2 and 3). Here, we define structural agreement as agreement with predefined formatting constraints. In a majority of cases, conformance was described as conformance, consistency, or representation and implied the use of some predefined structure, value, or format. Structural agreement most often depended on the usage of a correct data type and unit if necessary.

Eleven (11%) papers assessed bias most commonly using data element agreement (Table 2). We define bias as a dimension of data quality as missingness not at random. For example, some authors identified the pattern that sicker patients have higher levels of data completeness which implies that exclusion based on complete records will select a biased sample in terms of patient health levels.[31,77,88,89] Additionally, some authors highlighted the differences in data availability from structured versus unstructured data and suggested the bias resulting from using only one of the forms of EHR data.[73,77] Differential recording of patient attributes by race also constituted an example of bias.[72] Although sometimes similar to the dimension of completeness in considering

missing data, it can be seen from these examples that the dimension of bias further examines missing data rates in the presence of other variables.

Tools

Tools included in this review were often described as frameworks or ontologies for assessing EHR data quality. Eighteen of the 19 tools were built to assess structured data, while only 2 of the tools were built to assess unstructured data. Tools most often assessed completeness (86%), plausibility (50%), and correctness (45%), though currency, concordance, and conformance were also considered.

From the iterative review process, we established 4 themes in the tool development papers. The first theme addresses the task or project dependency of DQA.[12,22,25,32,84,90–92] These tools provide a mechanism to adjust aspects of the DQA based on the project or use case at hand. However, this task dependence must be balanced with the second theme, which highlights the necessity for scalable tool development.[11,12,22,32,85,91–96] These authors highlight the fact that consistent and comparable DQA requires identical approaches across domains. Due to the immense range of domain applications, this comparability is challenging to achieve.

The other 2 themes provide suggestions for improving a streamlined DQA. The third theme advocates more consistent use of a CDM.[11,25,69,75,84,85,91,92] A CDM could facilitate the scalability of a tool and allow for easier comparison of data quality across different domains. The fourth theme recommends the automation of DQA.[11,12,25,54,75,85,92,95–97] Initial attempts to automate DQA included software packages that can be applied to various EHR datasets,[85,94] applied to EHR data in a specific type of system,[92] and rule lists or frameworks that can be assessed on different EHR datasets.[12,17,32] Automation would support both the timeliness of DQA and the ability to use a single DQA tool across multiple domains.

Opinion pieces

Opinion pieces often consisted of a collection of ideas derived from expert panels or stakeholders through surveys and interviews.[59,98,99] Due to the process of collecting different opinions, the opinion pieces developed the general notion that a collaborative team is necessary and helpful for developing DQA.[36,99,100] These opinion pieces echoed many of the themes highlighted in the tool category. Most notably, they agreed that CDMs would be useful both for completing and comparing DQA methods.[36,53,59,98,101] However, they also acknowledge the task dependency of DQA as a limiting factor

Table 3. Dimensions of data quality and methods of assessment

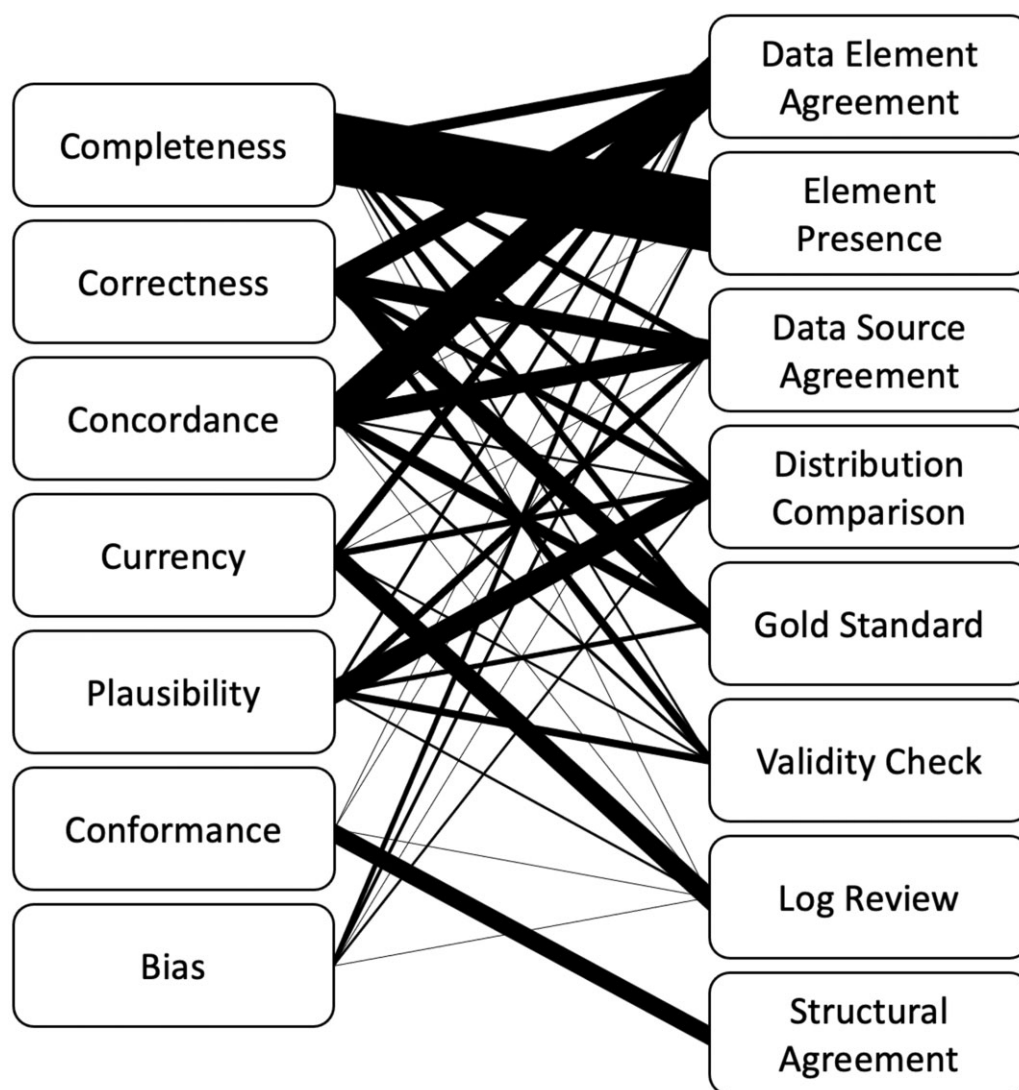| Dimension	Completeness	Correctness	Concordance	Currency	Plausibility	Conformance	Bias	Total
Method								
Data element agreement	10[17–19,21,23,24,27,28,70,77]	20[18,19,29–32,34–38,40–42,48,51,60,63,67,77]	35[12,17,18,25,30,31,34,38,41,42,48,50,52–55,60,63–66,68–74,76–80,83,102]	8[22,35,40,42,71,73,80,103]	3[11,69,96]	1[11]	5[31,72,73,77,104]	82
Element presence	64[11,12,17,18,20,21,23,24,26,27,29,32,35,36,38,39,42,45,47,49,52–56,59–63,66,67,69,70,73,77,79–90,92,94,96,99,102,104–114]	0	1[17]	0	0	1[11]	3[27,88,104]	69
Data source agreement	7[18,20,22,25,33,57,79]	17[18,25,33,35,38,39,48,53–55,58,59,61,62,75,97,115]	17[17,18,25,31,53,54,56,57,60,65,66,74,80–83,87]	1[79]	7[18,22,63,65,81,86,116]	0	1[57]	50
Distribution comparison	6[22,23,33,42,89,112]	8[18,23,32,59,63,67,85,107]	2[63,87]	6[18,42,53,60,93,101]	20[12,18,20,22,42,53,55,59,62,65,79,83,85,86,92,94,95,106,116,117]	0	2[89,103]	44
Gold standard	6[19,22,26,68,79,84]	18[19,33,34,40–47,59–62,68,97,115]	11[42,46,60,69,79,81,83,85–87,102]	0	5[42,69,81,86,93]	0	0	40
Validity check	2[42,49]	8[22,29,49,50,56,58,63,67]	3[63,80,81]	2[22,42]	7[20,42,56,83,92,94,96]	0	0	22
Log review	1[17]	0	1[17]	18[12,18,23,25,32,49,53–55,60,65,71,73,79,101,104,111,113]	2[11,62]	1[11]	1[111]	24
Structural xxsagreement	0	0	0	0	0	16[11,20,25,26,47,53,54,84,85,90,94,96,101,106,113,117]	0	16
Total	76	53	46	35	29	18	11	

Figure 2. Map comparing dimensions of data quality and methods used to assess dimensions of data quality. Dimensions are listed in the boxes on the left and methods are listed in the boxes on the right. The weight of an edge indicates the frequency of that combination. This figure presents an updated version of Figure 1 from the original review.[8]

for comparison of qualities between different assessments at this point in time.[18,59,99,100]

DISCUSSION

EHR data quality is of paramount importance as EHR data continues to be increasingly leveraged for biomedical research purposes. In order to understand trends in DQA of EHR data, we extended a 2013 literature review[8] on the topic to present day. In 2013, Weiskopf established 5 dimensions of EHR data quality and 7 methods by which to assess these dimensions. Since 2013, we have found a general increase in the number of dimensions assessed per paper and the number of methods used along with the addition of 2 dimensions to the framework.

In regards to *a priori* specified dimensions, we found an increase in the proportion of papers that assess completeness, concordance, plausibility, and currency, and a decrease in the proportion of papers that assess correctness as they relate to EHR data since 2013. This decrease in proportion of papers

assessing correctness should not be taken at face value as concordance and plausibility can sometimes be considered a subset of other dimensions.[8] Similarly, there was an increase in the proportion of papers in which each method was used. However, data element agreement, element presence, and data source agreement surpassed the use of a gold standard as the most common methods in that order. The trend towards using methods other than a gold standard comparison is positive as there are noted challenges that go along with establishing a gold standard comparator.[8]

In addition to Weiskopf's original 5 dimensions of data quality, we propose conformance and bias as additional and meaningful dimensions of data quality. Our definition of conformant data, or data that complies with a predefined relational structure, aligns with the definition of value conformance from Kahn et al[9] and extends Weiskopf's model of DQA to include aspects of Wang and Strong's representational data quality.[7] This represents a shift in DQA practices since the 2013 review which only identified dimensions that focused on intrinsic and contextual DQA, a shift that may be

partly due to the adoption of frameworks like the one presented by Kahn et al[9] and the increased adoption of CDMs both within research networks and collaborative studies.[11,53] Structured data lend themselves well to intrinsic and contextual approaches as they have likely been assessed for conformance, while unstructured data in the EHR or an associated source, such as clinical registries, may still require assessment of conformance.

The second dimension we added to Weiskopf's model of DQA was bias, or missingness not at random, which is often due to information or measurement bias.[118] Generally in informatics research, there are different levels of missingness, some of which can be ignored in secondary analyses and some of which impact the outcome of the analyses.[119] Bias is one mechanism for understanding whether or not missingness in EHR data is ignorable and is therefore increasingly important to consider as it has direct implications for downstream research. Recently, there have been many projects in the health informatics domain which highlight the damaging effects of biased data on research outcomes. Bias in EHR data can cause bias in the machine learning and artificial intelligence models developed using the data.[120,121] These biased results may then have a negative impact on patient care when machine learning models are used in decision support tools in clinical practice.[122] Patients may be assigned incorrect risk scores or be given incorrect treatment recommendations.[123] This is especially problematic when biased results, in turn, perpetuate systematic inequities in healthcare systems and delivery at the individual and population levels.

Bias as a dimension of EHR data quality also provides an example of an underlying mechanism behind general data quality issues. When data quality issues do occur, they likely can be attributed to some underlying mechanism. These mechanisms range from data entry or documentation errors to larger problems within EHR storage or warehouse software to institutional level barriers to accessing care. Understanding from where a data quality error may stem can help to identify key points in the data lifecycle at which to assess and improve data quality. Such an observation also argues for the recognition of the dynamic, complex systems that influence or impact data quality in "real world" settings.

Despite the consistent patterns of DQA in the literature found by this review, researchers largely developed DQA on a project-by-project basis. The methods used to assess data quality were repeatedly implemented across many applications although assessing a consistent collection of dimensions. The repetitive patterns of DQA are not practical in terms of time and resources in our current research environment as EHR data continues to be commonly used for downstream analysis. For this reason, we recommend and highlight the emerging theme of DQA automation as discussed in many of the opinion pieces and tools. This review emphasized the movement towards automating DQA in the development of DQA tools. Examples of automation include software packages, rule lists, and frameworks.

The prospect of automating DQA also requires interrogation of the data lifecycle to determine optimal points at which to assess data quality. Although the majority of DQA in the literature occurs after extraction from the EHR, some of the proposed tools transitioned to DQA at an EHR software level.[92] Based on prior work and experience using EHR data for downstream analysis, we identified the original entry point of data into an EHR, the transition to a data warehouse, or after extraction for a specific project as natural opportunities for DQA. It may be the case that certain dimensions of DQA are available at different stages in the data lifecycle. Implementation of automated DQA checks across an EHR ecosystem could help improve interoperability and further the transition to a comparable model of DQA.

In addition to automation, we should give further consideration to the balance between a scalable tool and a task specific tool. As data requirements differ between systems and projects, we will need a flexible tool in order to be able to assess data quality across many applications. One potential solution to this problem is the usage of a CDM to support interoperability and enable the development of reusable DQA tools.[25,36,53,59,69,75,84,85,91,92,98,101]

Limitations

There are a few limitations of this review to consider. First, the paper selection process was subjective as it was only performed by one author. For this reason, authors and other reviewers may not agree with our classifications. In addition, we were unable to review all of the initial results due to resource constraints. Our screening process could be considered a convenience sample which optimizes for recent research and highly visible research based on citation frequency. It is possible that selecting literature in order of citation frequency may identify papers cited for clinical research content rather than data quality content. However, the primary objective of all included papers was DQA, so we believe the number of citations implies a larger visibility for the DQA methods in the research community regardless of the citation purpose. A future review could take the time to review all initial search results rather than adopting our dual importance and emergence approach or could ensure that selection by citation frequency optimizes for data quality literature.

CONCLUSION

Although high quality EHR data are necessary to support patient care and secondary analyses, there do not exist standard methods for assessing EHR data quality. We extended a 2013 literature review on EHR DQA to evaluate changes and improvements in DQA approaches. There has been an increase in the number of dimensions of DQA and the methods by which to assess the dimensions of EHR data quality in recent years. However, there still does not exist a standard approach for DQA of EHR data, so future work should focus on the development of DQA tools and potential automation of such tools.

FUNDING

AUTHOR CONTRIBUTIONS

AEL, PROP, and AG conceived and designed the study. AEL completed the literature review and analysis, drafted and revised the manuscript, and prepared tables and figures. AEL, NW, PROP, and AG participated in the literature review interpretation and drafted the manuscript. All authors reviewed and revised the manuscript, and approved the final version for submission.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of the American Medical Informatics Association* online.

CONFLICT OF INTEREST STATEMENT

None declared.

DATA AVAILABILITY

The data underlying this article are available in the article and in its online supplementary material.

REFERENCES

1. Nordo AH, Levaux HP, Becnel LB, *et al.* Use of EHRs data for clinical research: historical progress and current applications. *Learn Health Syst* 2019; 3 (1): e10076.
2. Maddox TM, Rumsfeld JS, Payne PRO. Questions for artificial intelligence in health care. *JAMA* 2019; 321 (1): 31–2.
3. Beresniak A, Schmidt A, Proeve J, *et al.* Cost-benefit assessment of using electronic health records data for clinical research versus current practices: contribution of the Electronic Health Records for Clinical Research (EHR4CR) European Project. *Contemp Clin Trials* 2016; 46: 85–91.
4. Dagliati A, Malovini A, Tibollo V, Bellazzi R. Health informatics and EHR to support clinical research in the COVID-19 pandemic: an overview. *Brief Bioinform* 2021; 22 (2): 812–22.
5. Xu H, Buckeridge DL, Wang F, Tarczy-Hornoch P. Novel informatics approaches to COVID-19 research: from methods to applications. *J Biomed Inform* 2022; 129: 104028.
6. Edmondson ME, Reimer AP. Challenges frequently encountered in the secondary use of electronic medical record data for research. *Comput Inform Nurs* 2020; 38 (7): 338–48.
7. Wang RY, Strong DM. Beyond accuracy: what data quality means to data consumers. *J Manag Inform Syst* 1996; 12 (4): 5–33.
8. Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. *J Am Med Inform Assoc* 2013; 20 (1): 144–51.
9. 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 (Wash DC)* 2016; 4 (1): 1244.
10. Qualls LG, Phillips TA, Hammill BG, *et al.* Evaluating foundational data quality in the national patient-centered clinical research network (PCORnet®). *EGEMS (Wash DC)* 2018; 6 (1): 3.
11. Engel N, Wang H, Jiang X, *et al.* EHR data quality assessment tools and issue reporting workflows for the 'All of Us' research program clinical data research network. *AMIA Annu Symp Proc* 2022; 2022: 186–95.
12. Wang Z, Talburt JR, Wu N, Dagtas S, Zozus MN. A rule-based data quality assessment system for electronic health record data. *Appl Clin Inform* 2020; 11 (4): 622–34.
13. Hripcsak G, Duke JD, Shah NH, *et al.* Observational Health Data Sciences and Informatics (OHDSI): opportunities for observational researchers. *Stud Health Technol Inform* 2015; 216: 574–8.
14. Huser V, DeFalco FJ, Schuemie M, *et al.* Multisite evaluation of a data quality tool for patient-level clinical data sets. *EGEMS (Wash DC)* 2016; 4 (1): 1239.
15. Pfaff ER, Girvin AT, Gabriel DL, *et al.*; N3C Consortium. Synergies between centralized and federated approaches to data quality: a report from the national COVID cohort collaborative. *J Am Med Inform Assoc* 2022; 29 (4): 609–18.
16. Page MK, McKenzie JE, Boutron I, *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021; 372: n71. doi:10.1136/bmj.n71.
17. Reimer AP, Milinovich A, Madigan EA. Data quality assessment framework to assess electronic medical record data for use in research. *Int J Med Inform* 2016; 90: 40–7.
18. Feder SL. Data quality in electronic health records research: quality domains and assessment methods. *West J Nurs Res* 2018; 40 (5): 753–66.
19. Hernandez-Boussard T, Monda KL, Crespo BC, Riskin D. Real world evidence in cardiovascular medicine: ensuring data validity in electronic health record-based studies. *J Am Med Inform Assoc* 2019; 26 (11): 1189–94.
20. Khare R, Utidjian L, Ruth BJ, *et al.* A longitudinal analysis of data quality in a large pediatric data research network. *J Am Med Inform Assoc* 2017; 24 (6): 1072–9.
21. Singer A, Yakubovich S, Kroeker AL, Dufault B, Duarte R, Katz A. Data quality of electronic medical records in Manitoba: do problem lists accurately reflect chronic disease billing diagnoses? *J Am Med Inform Assoc* 2016; 23 (6): 1107–12.
22. Terry AL, Stewart M, Cejic S, *et al.* A basic model for assessing primary health care electronic medical record data quality. *BMC Med Inform Decis Mak* 2019; 19 (1): 30.
23. 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. *AMIA Annu Symp Proc* 2022; 2022: 293–302.
24. Dixon BE, Siegel JA, Oemig TV, Grannis SJ. Electronic health information quality challenges and interventions to improve public health surveillance data and practice. *Public Health Rep* 2013; 128 (6): 546–53.
25. Johnson SG, Speedie S, Simon G, Kumar V, Westra BL. A data quality ontology for the secondary use of EHR data. *AMIA Annu Symp Proc* 2015; 2015: 1937–46.
26. Sirgo G, Esteban F, Gómez J, *et al.* Validation of the ICU-DaMa tool for automatically extracting variables for minimum dataset and quality indicators: The importance of data quality assessment. *Int J Med Inform* 2018; 112: 166–72.
27. Samal L, Linder JA, Bates DW, Wright A. Electronic problem list documentation of chronic kidney disease and quality of care. *BMC Nephrol* 2014; 15: 70.
28. Wright A, McCoy AB, Hickman TT, *et al.* Problem list completeness in electronic health records: a multi-site study and assessment of success factors. *Int J Med Inform* 2015; 84 (10): 784–90.
29. Puttkammer N, Baseman JG, Devine EB, *et al.* An assessment of data quality in a multi-site electronic medical record system in Haiti. *Int J Med Inform* 2016; 86: 104–16.
30. Mays JA, Mathias PC. Measuring the rate of manual transcription error in outpatient point-of-care testing. *J Am Med Inform Assoc* 2019; 26 (3): 269–72.
31. Nicholson BD, Aveyard P, Hamilton W, *et al.* The internal validation of weight and weight change coding using weight measurement data within the UK primary care electronic health record. *Clin Epidemiol* 2019; 11: 145–55.
32. Weiskopf NG, Bakken S, Hripcsak G, Weng C. A data quality assessment guideline for electronic health record data reuse. *EGEMS (Wash DC)* 2017; 5 (1): 14.
33. Liaw ST, Taggart J, Yu H, de Lusignan S, Kuziemsky C, Hayen A. Integrating electronic health record information to support integrated care: practical application of ontologies to improve the accuracy of diabetes disease registers. *J Biomed Inform* 2014; 52: 364–72.
34. Daskivich TJ, Abedi G, Kaplan SH, *et al.* Electronic health record problem lists: accurate enough for risk adjustment? *Am J Manag Care* 2018; 24 (1): e24–9.
35. 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): 434–40.

36. Kohane IS, Aronow BJ, Avillach P, *et al.*; Consortium for Clinical Characterization of COVID-19 by EHR (4CE). What every reader should know about studies using electronic health record data but may be afraid to ask. *J Med Internet Res* 2021; 23 (3): e22219.

37. Diaz-Garelli JF, Strowd R, Ahmed T, *et al.* A tale of three subspecialties: diagnosis recording patterns are internally consistent but specialty-dependent. *JAMIA Open* 2019; 2 (3): 369–77.

38. Dentler K, Cornet R, ten Teije A, *et al.* Influence of data quality on computed Dutch hospital quality indicators: a case study in colorectal cancer surgery. *BMC Med Inform Decis Mak* 2014; 14: 32.

39. Scholte M, van Dulmen SA, Neeleman-Van der Steen CW, van der Wees PJ, Nijhuis-van der Sanden MW, Braspenning J. Data extraction from electronic health records (EHRs) for quality measurement of the physical therapy process: comparison between EHR data and survey data. *BMC Med Inform Decis Mak* 2016; 16 (1): 141.

40. Diaz-Garelli JF, Strowd R, Wells BJ, Ahmed T, Merrill R, Topaloglu U. Lost in translation: diagnosis records show more inaccuracies after biopsy in oncology care EHRs. *AMIA Jt Summits Transl Sci Proc* 2019; 2019: 325–34.

41. Coleman N, Halas G, Peeler W, Casaclang N, Williamson T, Katz A. From patient care to research: a validation study examining the factors contributing to data quality in a primary care electronic medical record database. *BMC Fam Pract* 2015; 16: 11.

42. Verma AA, Pasricha SV, Jung HY, *et al.* Assessing the quality of clinical and administrative data extracted from hospitals: the General Medicine Inpatient Initiative (GEMINI) experience. *J Am Med Inform Assoc* 2021; 28 (3): 578–87.

43. Martin S, Wagner J, Lupulescu-Mann N, *et al.* Comparison of EHR-based diagnosis documentation locations to a gold standard for risk stratification in patients with multiple chronic conditions. *Appl Clin Inform* 2017; 8 (3): 794–809.

44. Kreuzthaler M, Schulz S, Berghold A. Secondary use of electronic health records for building cohort studies through top-down information extraction. *J Biomed Inform* 2015; 53: 188–95.

45. Diaz-Garelli F, Strowd R, Lawson VL, *et al.* Workflow differences affect data accuracy in oncologic EHRs: a first step toward detangling the diagnosis data Babel. *JCO Clin Cancer Inform* 2020; 4: 529–38.

46. Xi N, Wallace R, Agarwal G, Chan D, Gupta GA. S. Identifying patients with asthma in primary care electronic medical record systems chart analysis-based electronic algorithm validation study. *Can Fam Phys* 2015; 61 (10): e474–83.

47. Rahimi A, Liaw ST, Taggart J, Ray P, Yu H. Validating an ontology-based algorithm to identify patients with type 2 diabetes mellitus in electronic health records. *Int J Med Inform* 2014; 83 (10): 768–78.

48. Knake LA, Ahuja M, McDonald EL, *et al.* Quality of EHR data extractions for studies of preterm birth in a tertiary care center: guidelines for obtaining reliable data. *BMC Pediatr* 2016; 16: 59.

49. Yadav S, Kazanji N, K CN, *et al.* Comparison of accuracy of physical examination findings in initial progress notes between paper charts and a newly implemented electronic health record. *J Am Med Inform Assoc* 2017; 24 (1): 140–4.

50. Binkheder S, Asiri MA, Altowayan KW, *et al.* Real-world evidence of COVID-19 patients' data quality in the electronic health records. *Healthcare (Basel)* 2021; 9 (12): 1648.

51. Ammann EM, Kalsekar I, Yoo A, *et al.* Assessment of obesity prevalence and validity of obesity diagnoses coded in claims data for selected surgical populations: a retrospective, observational study. *Medicine (Baltimore)* 2019; 98 (29): e16438.

52. Angier H, Gold R, Gallia C, *et al.* Variation in outcomes of quality measurement by data source. *Pediatrics* 2014; 133 (6): e1676–82.

53. Brown JS, Kahn M, Toh S. Data quality assessment for comparative effectiveness research in distributed data networks. *Med Care* 2013; 51 (8 Suppl 3): S22–9.

54. 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 (Wash DC)* 2016; 4 (1): 1201.

55. Groenhof TKJ, Koers LR, Blasse E, *et al.*; UCC-CVRM Study Groups. Data mining information from electronic health records produced high yield and accuracy for current smoking status. *J Clin Epidemiol* 2020; 118: 100–6.

56. Barkhuysen P, de Grauw W, Akkermans R, Donkers J, Schers H, Biermans M. Is the quality of data in an electronic medical record sufficient for assessing the quality of primary care? *J Am Med Inform Assoc* 2014; 21 (4): 692–8.

57. Thompson CA, Jin A, Luft HS, *et al.* Population-based registry linkages to improve validity of electronic health record-based cancer research. *Cancer Epidemiol Biomarkers Prev* 2020; 29 (4): 796–806.

58. Sollie A, Sijmons RH, Helsper C, Numans ME. Reusability of coded data in the primary care electronic medical record: a dynamic cohort study concerning cancer diagnoses. *Int J Med Inform* 2017; 99: 45–52.

59. Richesson RL, Hammond WE, Nahm M, *et al.* Electronic health records based phenotyping in next-generation clinical trials: a perspective from the NIH Health Care Systems Collaboratory. *J Am Med Inform Assoc* 2013; 20 (e2): e226–31.

60. Polubriaginof FCG, Ryan P, Salmasian H, *et al.* Challenges with quality of race and ethnicity data in observational databases. *J Am Med Inform Assoc* 2019; 26 (8–9): 730–6.

61. Pedrera-Jimenez M, Garcia-Barrio N, Rubio-Mayo P, *et al.* Making EHRs trustable: a quality analysis of EHR-derived datasets for COVID-19 research. *Stud Health Technol Inform* 2022; 294: 164–8.

62. Thuraisingam S, Chondros P, Dowsey MM, *et al.* Assessing the suitability of general practice electronic health records for clinical prediction model development: a data quality assessment. *BMC Med Inform Decis Mak* 2021; 21 (1): 297.

63. Polubriaginof F, Salmasian H, Albert DA, Vawdrey DK. Challenges with collecting smoking status in electronic health records. *AMIA Annu Symp Proc* 2017; 2017: 1392–400.

64. Just BH, Marc D, Munns M, Sandefer R. Why patient matching is a challenge: research on master patient index (MPI) data discrepancies in key identifying fields. *Perspect Health Inf Manag* 2016; 13 (Spring): 1e.

65. Horsfall L, Walters K, Petersen I. Identifying periods of acceptable computer usage in primary care research databases. *Pharmacoepidemiol Drug Saf* 2013; 22 (1): 64–9.

66. Lynch KE, Deppen SA, DuVall SL, *et al.* Incrementally transforming electronic medical records into the observational medical outcomes partnership common data model: a multidimensional quality assurance approach. *Appl Clin Inform* 2019; 10 (5): 794–803.

67. Carsley S, Birken CS, Parkin P, Pullenayegum E, Tu K. Completeness and accuracy of anthropometric measurements in electronic medical records for children attending primary care. *J Innov Health Inform* 2018; 25 (1): 963.

68. Ostropolets A, Reich C, Ryan P, Shang N, Hripcsak G, Weng C. Adapting electronic health records-derived phenotypes to claims data: lessons learned in using limited clinical data for phenotyping. *J Biomed Inform* 2020; 102: 103363.

69. Cusick MM, Sholle ET, Davila MA, Kabariti J, Cole CL, Campion TR Jr. A method to improve availability and quality of patient race data in an electronic health record system. *Appl Clin Inform* 2020; 11 (5): 785–91.

70. Chen ES, Carter EW, Sarkar IN, Winden TJ, Melton GB. Examining the use, contents, and quality of free-text tobacco use documentation in the electronic health record. *AMIA Annu Symp Proc* 2014; 2014: 366–74.

71. Ford E, Carroll J, Smith H, *et al.* What evidence is there for a delay in diagnostic coding of RA in UK general practice records? An observational study of free text. *BMJ Open* 2016; 6 (6): e010393.

72. Sholle ET, Pinheiro LC, Adekkanattu P, *et al*. Underserved populations with missing race ethnicity data differ significantly from those with structured race/ethnicity documentation. *J Am Med Inform Assoc* 2019; 26 (8–9): 722–9.

73. Dumas RP, Chreiman KM, Seamon MJ, *et al*. Benchmarking emergency department thoracotomy: using trauma video review to generate procedural norms. *Injury* 2018; 49 (9): 1687–92.

74. Brundin-Mather R, Soo A, Zuege DJ, *et al*. Secondary EMR data for quality improvement and research: A comparison of manual and electronic data collection from an integrated critical care electronic medical record system. *J Crit Care* 2018; 47: 295–301.

75. Parr SK, Shotwell MS, Jeffery AD, Lasko TA, Matheny ME. Automated mapping of laboratory tests to LOINC codes using noisy labels in a national electronic health record system database. *J Am Med Inform Assoc* 2018; 25 (10): 1292–300.

76. Weiskopf NG, Cohen AM, Hannan J, Jarmon T, Dorr DA. Towards augmenting structured EHR data: a comparison of manual chart review and patient self-report. *AMIA Annu Symp Proc* 2019; 2019: 903–12.

77. Liede A, Hernandez RK, Roth M, Calkins G, Larrabee K, Nicacio L. Validation of International Classification of Diseases coding for bone metastases in electronic health records using technology-enabled abstraction. *Clin Epidemiol* 2015; 7: 441–8.

78. Rhee C, Jentzsch MS, Kadri SS, *et al*.; Centers for Disease Control and Prevention (CDC) Prevention Epicenters Program. Variation in identifying sepsis and organ dysfunction using administrative versus electronic clinical data and impact on hospital outcome comparisons. *Crit Care Med* 2019; 47 (4): 493–500.

79. Ammann EM, Kalsekar I, Yoo A, Johnston SS. Validation of body mass index (BMI)-related ICD-9-CM and ICD-10-CM administrative diagnosis codes recorded in US claims data. *Pharmacoepidemiol Drug Saf* 2018; 27 (10): 1092–100.

80. Bailey LC, Milov DE, Kelleher K, *et al*. Multi-institutional sharing of electronic health record data to assess childhood obesity. *PLoS One* 2013; 8 (6): e66192.

81. McVeigh KH, Newton-Dame R, Chan PY, *et al*. Can electronic health records be used for population health surveillance? Validating population health metrics against established survey data. *EGEMS (Wash DC)* 2016; 4 (1): 1267.

82. Lee SJ, Grobe JE, Tiro JA. Assessing race and ethnicity data quality across cancer registries and EMRs in two hospitals. *J Am Med Inform Assoc* 2016; 23 (3): 627–34.

83. Horth RZ, Wagstaff S, Jeppson T, *et al*. Use of electronic health records from a statewide health information exchange to support public health surveillance of diabetes and hypertension. *BMC Public Health* 2019; 19 (1): 1106.

84. Lee K, Weiskopf N, Pathak J. A framework for data quality assessment in clinical research datasets. *AMIA Annu Symp Proc* 2017; 2017: 1080–9.

85. Pezoulas VC, Kourou KD, Kalatzis F, *et al*. Medical data quality assessment: on the development of an automated framework for medical data curation. *Comput Biol Med* 2019; 107: 270–83.

86. Funk LM, Shan Y, Voils CI, Kloke J, Hanrahan LP. Electronic health record data versus the National Health and Nutrition Examination Survey (NHANES): a comparison of overweight and obesity rates. *Med Care* 2017; 55 (6): 598–605.

87. Newton-Dame R, McVeigh KH, Schreibstein L, *et al*. Design of the New York City macroscope: innovations in population health surveillance using electronic health records. *EGEMS (Wash DC)* 2016; 4 (1): 1265.

88. Weiskopf NG, Rusanov A, Weng C. Sick patients have more data: the non-random completeness of electronic health records. *AMIA Annu Symp Proc* 2013; 2013: 1472–7.

89. Weber GM, Adams WG, Bernstam EV, *et al*. Biases introduced by filtering electronic health records for patients with "complete data". *J Am Med Inform Assoc* 2017; 24 (6): 1134–41.

90. Johnson SG, Speedie S, Simon G, Kumar V, Westra BL. Application of an ontology for characterizing data quality for a secondary use of EHR data. *Appl Clin Inform* 2016; 7 (1): 69–88.

91. van der Bij S, Khan N, Ten Veen P, de Bakker DH, Verheij RA. Improving the quality of EHR recording in primary care: a data quality feedback tool. *J Am Med Inform Assoc* 2017; 24 (1): 81–7.

92. Tute E, Scheffner I, Marschollek M. A method for interoperable knowledge-based data quality assessment. *BMC Med Inform Decis Mak* 2021; 21 (1): 93.

93. Estiri H, Vasey S, Murphy SN. Generative transfer learning for measuring plausibility of EHR diagnosis records. *J Am Med Inform Assoc* 2021; 28 (3): 559–68.

94. Kapsner LA, Kampf MO, Seuchter SA, *et al*. Moving towards an EHR data quality framework: the MIRACUM approach. *Stud Health Technol Inform* 2019; 267: 247–53.

95. Estiri H, Klann JG, Murphy SN. A clustering approach for detecting implausible observation values in electronic health records data. *BMC Med Inform Decis Mak* 2019; 19 (1): 142.

96. Mang JM, Seuchter SA, Gulden C, *et al*. DQAgui: a graphical user interface for the MIRACUM data quality assessment tool. *BMC Med Inform Decis Mak* 2022; 22 (1): 213.

97. Lai KH, Topaz M, Goss FR, Zhou L. Automated misspelling detection and correction in clinical free-text records. *J Biomed Inform* 2015; 55: 188–95.

98. Callahan T, Barnard J, Helmkamp L, Maertens J, Kahn M. Reporting data quality assessment results: identifying individual and organizational barriers and solutions. *EGEMS (Wash DC)* 2017; 5 (1): 16.

99. Dixon BE, Rosenman M, Xia Y, Grannis SJ. A vision for the systematic monitoring and improvement of the quality of electronic health data. *Stud Health Technol Inform* 2013; 192: 884–8.

100. Sudat SEK, Robinson SC, Mudiganti S, Mani A, Pressman AR. Mind the clinical-analytic gap: electronic health records and COVID-19 pandemic response. *J Biomed Inform* 2021; 116: 103715.

101. Colquhoun DA, Shanks AM, Kapeles SR, *et al*. Considerations for integration of perioperative electronic health records across institutions for research and quality improvement: the approach taken by the multicenter perioperative outcomes group. *Anesth Analg* 2020; 130 (5): 1133–46.

102. 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): e0195362.

103. Rockenschaub P, Nguyen V, Aldridge RW, Acosta D, García-Gómez JM, Sáez C. Data-driven discovery of changes in clinical code usage over time: a case-study on changes in cardiovascular disease recording in two English electronic health records databases (2001–2015). *BMJ Open* 2020; 10 (2): e034396.

104. Sidky H, Young JC, Girvin AT, *et al*.; The N3C Consortium. Data quality considerations for evaluating COVID-19 treatments using real world data: learnings from the National COVID Cohort Collaborative (N3C). *BMC Med Res Methodol* 2023; 23 (1): 46.

105. Weiskopf NG, Hripcsak G, Swaminathan S, Weng C. Defining and measuring completeness of electronic health records for secondary use. *J Biomed Inform* 2013; 46 (5): 830–6.

106. 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: 104804.

107. Taggart J, Liaw ST, Yu H. Structured data quality reports to improve EHR data quality. *Int J Med Inform* 2015; 84 (12): 1094–8.

108. 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): 349–54.

109. Lee WC, Veeranki SP, Serag H, Eschbach K, Smith KD. Improving the collection of race, ethnicity, and language data to reduce healthcare disparities: a case study from an academic medical center. *Perspect Health Inf Manag* 2016; 13 (Fall): 1g.

110. Garcia BH, Djønne BS, Skjold F, Mellingen EM, Aag TI. Quality of medication information in discharge summaries from hospitals:

an audit of electronic patient records. *Int J Clin Pharm* 2017; 39 (6): 1331–7.

111. Wiley KK, Mendonca E, Blackburn J, Menachemi N, Groot MD, Vest JR. Quantifying electronic health record data quality in telehealth and office-based diabetes care. *Appl Clin Inform* 2022; 13 (5): 1172–80.

112. Ruddle RA, Adnan M, Hall M. Using set visualisation to find and explain patterns of missing values: a case study with NHS hospital episode statistics data. *BMJ Open* 2022; 12 (11): e064887.

113. Fu S, Wen A, Schaeferle GM, *et al*. Assessment of data quality variability across two EHR systems through a case study of postsurgical complications. *AMIA Annu Symp Proc* 2022; 2022: 196–205.

114. Vass A, Reinecke I, Boeker M, Prokosch HU, Gulden C. Availability of structured data elements in electronic health records for supporting patient recruitment in clinical trials. *Stud Health Technol Inform* 2022; 290: 130–4.

115. Samalik JM, Goldberg CS, Modi ZJ, *et al*. Discrepancies in race and ethnicity in the electronic health record compared to self-report [published online ahead of print November 23, 2022]. *J Racial Ethn Health Disparities* 2022.

116. Daymont C, Ross ME, Russell Localio A, Fiks AG, Wasserman RC, Grundmeier RW. Automated identification of implausible values in growth data from pediatric electronic health records. *J Am Med Inform Assoc* 2017; 24 (6): 1080–7.

117. Goldstein JE, Guo X, Boland MV, Smith KE. Visual acuity: assessment of data quality and usability in an electronic health record system. *Ophthalmol Sci* 2023; 3 (1): 100215.

118. Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A. A survey on bias and fairness in machine learning. *ACM Comput Surv* 2022; 54 (6): 1–35.

119. Rubin DB. Inference and missing data. *Biometrika* 1976; 63 (3): 581–92.

120. Andaur Navarro CL, Damen JAA, Takada T, *et al*. Risk of bias in studies on prediction models developed using supervised machine learning techniques: systematic review. *BMJ* 2021; 375: n2281.

121. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential biases in machine learning algorithms using electronic health record data. *JAMA Intern Med* 2018; 178 (11): 1544–7.

122. Char DS, Shah NH, Magnus D. Implementing machine learning in health care – addressing ethical challenges. *N Engl J Med* 2018; 378 (11): 981–3.

123. Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K. Artificial intelligence, bias and clinical safety. *BMJ Qual Saf* 2019; 28 (3): 231–7.