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

Abigail E Lewis
Washington University School of Medicine in St. Louis

Nicole Weiskopf
Oregon Health & Science University

Zachary B Abrams
Washington University School of Medicine in St. Louis

Randi Foraker
Washington University School of Medicine in St. Louis

Albert M Lai
Washington University School of Medicine in St. Louis

See next page for additional authors

Follow this and additional works at: https://digitalcommons.wustl.edu/oa_4

Part of the Medicine and Health Sciences Commons

Please let us know how this document benefits you.

Recommended Citation
https://digitalcommons.wustl.edu/oa_4/2279

This Open Access Publication is brought to you for free and open access by the Open Access Publications at Digital Commons@Becker. It has been accepted for inclusion in 2020-Current year OA Pubs by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.
Electronic health record data quality assessment and tools: a systematic review

Abigail E. Lewis1,2, Nicole Weiskopf3, Zachary B. Abrams10, Randi Foraker9, Albert M. Lai6, Philip R.O. Payne5, and Aditi Gupta2,*

1Division of Computational and Data Sciences, Washington University in St. Louis, St. Louis, Missouri, USA
2Institute for Informatics, Data Science and Biostatistics, Washington University in St. Louis, St. Louis, Missouri, USA
3Department 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. 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. Additionally, EHR data offer potential cost and time savings as an alternative to the primary collection of medical data for research purposes. 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. In addition to the potential benefits of using EHR data for research, however, there are also challenges, including ensuring data quality concerns, timely access, patient protections and confidentiality, and the ability to generalize results based on EHR data.

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. 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). 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. 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 al12 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 Largescal 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 review9 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:


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

This article is published in journal, volume, and issue date. To access the full text, please visit the journal’s website at [link].
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 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, an alternative data source, billing data, or physician agreement. 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, manual review by a physician, billing data, unstructured data, or an external data source. Many terms were used to
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,\textsuperscript{17,18,25,30,31,33,46,50,53,63–73} unstructured data,\textsuperscript{34,55,70–73} or an external data source.\textsuperscript{11,12,25,69,75,84,85,91,92} A CDM could facilitate the streamlined DQA. The third theme advocates more consistent and comparable DQA requires identical approaches across domains. Due to the immense range of domain applications, this comparability is challenging to achieve.

Opinion pieces

Opinion pieces often consisted of a collection of ideas derived from expert panels or stakeholders through surveys and interviews.\textsuperscript{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.\textsuperscript{12,22,32,84,90–92} 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.\textsuperscript{36,53,59,98,101} However, they also acknowledge the task dependency of DQA as a limiting factor.

### Table 2. Dimensions of data quality

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

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.\textsuperscript{12,22,25,32,84,90–92} These tools provide a mechanism to adjust aspects of the DQA tool.\textsuperscript{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.\textsuperscript{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.\textsuperscript{11,12,22,32,85,92,95–97} Initial attempts to automate DQA included software packages that can be applied to various EHR datasets.\textsuperscript{85,94} Applied to EHR data in a specific type of system,\textsuperscript{92} and rule lists or frameworks that can be assessed on different EHR datasets.\textsuperscript{12,17,32} Automation would support both the timeliness of DQA and the ability to use a single DQA tool across multiple domains.

### 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.

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.\textsuperscript{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.\textsuperscript{73,77} Differential recording of patient attributes by race also constituted an example of bias.\textsuperscript{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.
Table 3. Dimensions of data quality and methods of assessment

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Completeness</th>
<th>Correctness</th>
<th>Concordance</th>
<th>Currency</th>
<th>Plausibility</th>
<th>Conformance</th>
<th>Bias</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data element agreement</td>
<td>10</td>
<td>20</td>
<td>35</td>
<td>8</td>
<td>31</td>
<td>11</td>
<td>5</td>
<td>82</td>
</tr>
<tr>
<td>Element presence</td>
<td>64</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>11</td>
<td>3</td>
<td>82</td>
<td>69</td>
</tr>
<tr>
<td>Data source agreement</td>
<td>7</td>
<td>17</td>
<td>17</td>
<td>179</td>
<td>718</td>
<td>0</td>
<td>157</td>
<td>50</td>
</tr>
<tr>
<td>Distribution comparison</td>
<td>6</td>
<td>6</td>
<td>2</td>
<td>263</td>
<td>81</td>
<td>0</td>
<td>0</td>
<td>44</td>
</tr>
<tr>
<td>Gold standard</td>
<td>6</td>
<td>18</td>
<td>11</td>
<td>114</td>
<td>42</td>
<td>0</td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td>Validity check</td>
<td>2</td>
<td>17</td>
<td>12</td>
<td>222</td>
<td>2</td>
<td>111</td>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td>Log review</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>Structural xxsagreement</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>76</td>
<td>53</td>
<td>46</td>
<td>35</td>
<td>29</td>
<td>18</td>
<td>11</td>
<td>1734</td>
</tr>
</tbody>
</table>
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 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. 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.

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 and extends Weiskopf's model of DQA to include aspects of Wang and Strong's representational data quality. 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. and the increased adoption of CDMs both within research networks and collaborative studies. 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. 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. 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. 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. Patients may be assigned incorrect risk scores or be given incorrect treatment recommendations. 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. 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.

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
This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

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


