Check for
updates

Comparing Medical Record Abstraction (MRA) error rates in an observational study to pooled rates identified in the data quality literature

Maryam Y. Garza[1,2,9*], Tremaine B. Williams[1], Songthip Ounpraseuth[3], Zhuopei Hu[3], Jeannette Lee[3], Jessica Snowden[3,4], Anita C. Walden[5], Alan E. Simon[6], Lori A. Devlin[7], Leslie W. Young[8] and Meredith N. Zozus[2]

Abstract

Background Medical record abstraction (MRA) is a commonly used method for data collection in clinical research, but is prone to error, and the influence of quality control (QC) measures is seldom and inconsistently assessed during the course of a study. We employed a novel, standardized MRA-QC framework as part of an ongoing observational study in an effort to control MRA error rates. In order to assess the effectiveness of our framework, we compared our error rates against traditional MRA studies that had not reported using formalized MRA-QC methods. Thus, the objective of this study was to compare the MRA error rates derived from the literature with the error rates found in a study using MRA as the sole method of data collection that employed an MRA-QC framework.

Methods A comparison of the error rates derived from MRA-centric studies identified as part of a systematic literature review was conducted against those derived from an MRA-centric study that employed an MRA-QC framework to evaluate the effectiveness of the MRA-QC framework. An inverse variance-weighted meta-analytical method with Freeman-Tukey transformation was used to compute pooled effect size for both the MRA studies identified in the literature and the study that implemented the MRA-QC framework. The level of heterogeneity was assessed using the Q-statistic and Higgins and Thompson's I^2 statistic.

Results The overall error rate from the MRA literature was 6.57%. Error rates for the study using our MRA-QC framework were between 1.04% (optimistic, all-field rate) and 2.57% (conservative, populated-field rate), 4.00–5.53% points less than the observed rate from the literature ($p < 0.0001$).

*Correspondence:
Maryam Y. Garza
garzam46@uthscsa.edu

Full list of author information is available at the end of the article

Conclusions Review of the literature indicated that the accuracy associated with MRA varied widely across studies. However, our results demonstrate that, with appropriate training and continuous QC, MRA error rates can be significantly controlled during the course of a clinical research study.

Keywords Medical record abstraction, Data quality, Clinical research, Clinical data management, Data collection

Background

Computers have been used in clinical studies since the early 1960s, although initial attempts to integrate them into research workflows were experimental and sporadic [1, 2]. The early application of computers to health-related research spawned a plethora of methods for collecting and preparing data for analysis [3, 4]. These activities continue to be evaluated by metrics of cost, time, and quality [5]. While cost and time affect the feasibility of research, and timeliness is certainly critical to the conduct and oversight of research, the scientific validity of research conclusions depends on data accuracy [6].

Data accuracy is attributable to how data are collected, entered, and cleaned, or otherwise processed; and the assessment and quantification of data accuracy are crucial to scientific inquiry. Several texts describe approaches to general data quality management, independent of the domain area in which they are applied [7–10]. These works focus on general methods for assessing and documenting data quality, as well as methods for storing data in ways that maintain or improve their quality, but do not provide sufficient details on data collection and processing methods applicable to specific industries and types of data. Thus, they provide little to no guidance to investigators and research teams planning a clinical research endeavor or attempting to operationalize data collection and management.

In previous work, we extensively reviewed the clinical research data quality literature and identified gaps that necessitated a formal review and secondary analysis of this literature to characterize the data quality resulting from different data processing methods [11]. Through this effort, we quantified the average, overall error rates attributable to 4 major data processing methods used in clinical research (medical record abstraction [MRA]), optical scanning, single-data entry, and double-data entry) based on the data provided within 93 peer reviewed manuscripts [11]. Our results indicated that data quality varied widely by data processing method. Specifically, MRA was associated with error rates an order of magnitude greater than those associated with other data processing techniques (70–2,784 versus 2–650 errors per 10,000 fields, respectively) [11]. MRA was the most ubiquitous data processing method over the time period of the review. Unfortunately, in a more recent review of studies using MRA, the foundational quality assurance activities identified as important in MRA processes were rarely reported with clinical studies:

describing and stating the data source within the medical record 0 (0%); use of abstraction methods and tools 18 (50%); controlling the abstraction environment, such as to prevent interruption and distraction 0 (0%); and attention to abstraction human resources such as minimum qualifications and training 15 (42%) [12]. Only 3 (8.3%) of the articles reported measuring and controlling the MRA error rate [12].

In an effort to mitigate the inherent risks associated with MRA, we developed and employed a theory-based, quality control (QC) framework to support MRA activities in clinical research [13]. Our MRA-QC framework involved standardized MRA training prior to study implementation, as well as a continuous QC process carried out throughout the course of the study. We implemented and evaluated the MRA-QC framework within the context of the Advancing Clinical Trials in Neonatal Opioid Withdrawal Current Experience (ACT NOW CE) Study [14] to measure the influence of formalized MRA training and continuous QC on data quality. We then compared our findings with MRA error rates from the literature to better understand the potential influence of this framework on data quality in clinical research studies. Thus, the objective of this study was to compare the MRA error rates derived from the literature with the error rates found in a study using MRA as the sole method of data collection that employed an MRA-QC framework.

Methods

Preliminary work: Comprehensive literature review

As described in a separate manuscript, a systematic review of the literature was performed to "identify clinical research studies that evaluated the quality of data obtained from data processing methods typically used in clinical research" [11]. Here, we refer to this as the comprehensive literature review. The 93 manuscripts identified through the comprehensive literature review were categorized by data processing method (e.g., MRA, optical scanning, single-data entry, and double-data entry), and only those specific to MRA were considered for this meta-analysis.

Information retrieval: Meta-analysis of MRA error rates

Excluding manuscripts for which MRA was not the primary method of data collection from the comprehensive set yielded 64 MRA-centric manuscripts for inclusion in the meta-analysis (Fig. 1). For this evaluation, we
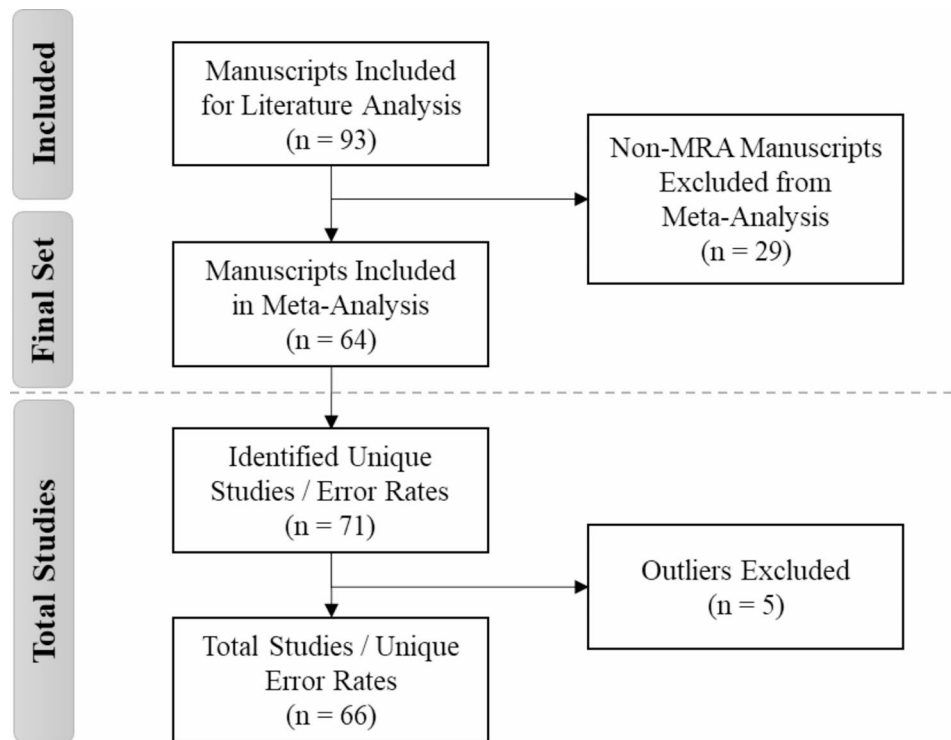
Fig. 1 Study selection flowchart: identification of MRA-centric literature for meta-analysis

referenced only this subset of MRA-centric studies to conduct a meta-analysis of the overall error rates as reported in the existing literature for comparison against the error rates derived for our study using the MRA-QC framework. Several manuscripts discussed multiple processing methods and/or studies, presenting unique error rates for each. Thus, within the set of 64 manuscripts, we identified 71 studies (or 71 unique error rates) for inclusion in the meta-analysis (see Additional File 1, Appendix A, Reference List A1 and Table A2).

Based on the residual and leave-one-out diagnostics [15–17], we identified 5 studies [18–22] that were deemed to be potential outliers. Thus, these studies were removed for the final meta-analysis to obtain the estimated error rate for the literature reviews. In order to derive an overall MRA error rate for comparison with our study, we performed a meta-analysis of single proportions to derive an overall error rate from the literature based on an inverse variance method [23, 24] and generalized linear mixed model approach using the R package "*metafor*" [25]. The general linear mixed model provided more robust estimates than traditional methods [26].

Comparison of MRA error rates to results of study using MRA-QC framework

Once the average MRA error rate across the literature was determined, we used that value to evaluate the effectiveness of a standardized MRA-QC framework

implemented as part of a retrospective research study for which MRA was the sole method for data collection. Briefly, the MRA-QC framework was implemented within the context of the ACT NOW CE Study[1] [27], a multicenter clinical research study sponsored by the National Institute of Health (NIH) through the Environmental Influences on Child Health Outcomes (ECHO) program [28]. Thirty IDeA[2] States Pediatric Clinical Trials Network (ISPCTN) [29, 30] and NICHD[3] Neonatal Research Network (NRN) [31] sites from across the U.S. participated in the study. Approximately 1,800 cases were abstracted across all study sites, of which a subset of cases (over 200) underwent a formalized QC process to identify data quality errors. Study coordinators at each site underwent a formalized MRA training program specific to the ACT NOW CE Study. As part of this training, the study investigators reviewed the full ACT NOW CE Study case report form with data collection staff at all

---

sites, walking through where in the EHR the data is typically charted, including potential secondary and tertiary locations in case the data was not available in the primary location. The primary objective was to standardize the way in which sites abstracted the data for the study for consistency and improved accuracy. Additional information on the ACT NOW CE Study [14], including details on the MRA training [32] and QC process [13], has been published elsewhere.

The overall error rates for the ACT NOW CE Study were compared to error rates from the literature. The overall error rate for the ACT NOW CE Study [13] was calculated using the same methodology used for calculating MRA error rates from the literature [11], based on the Society for Clinical Data Management's (SCDM) Good Clinical Data Management Practices (GCDMP) [33].

Error Rate Calculation Framework (Formula 1).

$$ErrorRate = \frac{Number\ of\ Errors\ Detected}{Number\ of\ Fields\ Collected} \quad (1)$$

All statistical tests were conducted at a two-sided significance level of 0.05. The analyses were performed using the R packages "*metafor*" and "*meta*". We employed an inverse variance-weighted meta-analytical method with Freeman-Tukey transformation to compute the pooled effect size [34] for both the MRA literature studies and the ACT NOW CE Study. During the analysis, records with studentized residuals exceeding an absolute value of 3 were identified as outliers and subsequently excluded, as this threshold suggests that the data point is unusual or influential in the model [35, 36]. The level of heterogeneity among studies was assessed using the Q-statistic and Higgins and Thompson's I^2 statistic [37]. For the MRA literature studies, we presented the error rates from the random effects model along with the corresponding 95% prediction intervals. Regarding the ACT NOW CE Study, we reported the all-field (optimistic) and populated-field (conservative) error rates (and the 95% confidence interval) from our study. A comparison was made between the error rates of the ACT NOW CE Study and the 95% prediction interval.

We decided to compare both optimistic and conservative rates to account for the variability in MRA error rates across the literature, in some cases simply due to the inconsistent methods used for counting the number of fields. As noted by Rostami and colleagues [38], a detailed examination of studies in the literature identified "significant methodological variation" in the way both errors and the total number of fields (error rate denominator) were counted. In some cases, the counting differences may cause "calculated error rates to vary by a factor of two or more" [38]. Accordingly, here, the all-field, optimistic rate was calculated using the larger, singular denominator ($N=312$), as the full set of data elements, regardless of data entry, was used. In comparison, the populated-field, conservative rate was calculated using a denominator that varied based only on the number of fields populated (non-null fields) in a case report form (N ranges from 1 to 312 depending on data entry). In other words, for the populated-field count, only fields for which data had been entered on the case report form (paper and/or electronic) were counted.

Results

The overall error rate from the literature meta-analysis (relying on data only from MRA-centric manuscripts) was 6.57%. The 95% prediction interval from the MRA literature ranged from 0.54 to 18.5%, providing a range within which we could anticipate the effects of future studies based on the existing literature. Notably, the error rates calculated from both the ACT NOW CE Study all-field and populated field calculations fell within this range (Table 1). Consequently, there was no statistical difference between the ACT NOW CE Study and the MRA literature.

Discussion

Through the pooled analysis of data error rates from the literature, we were able to establish an average, overall MRA error rate – approximately 6.57% (or 657 errors per 10,000 fields). We compared this rate (resulting from the MRA meta-analysis) to the rates calculated for the ACT NOW CE Study (1.04 – 2.57% or 104 to 257 errors per 10,000 fields) [13] and found that the error rates for the ACT NOW CE Study were substantially lower than those found in the peer reviewed literature – a difference of 553 (95% CI: 439, 667) per 10,000 fields (all-field total) and 400 (95% CI: 267, 533) per 10,000 fields (populated-field total). As cited in our previous work, we used both all-field and populated-field rates when calculating and presenting error rates for the ACT NOW CE Study, "to account for the variability in the calculation and reporting of error rates in the literature" [13, 38, 39]. Based on

Table 1 Error rate comparison: MRA Literature vs. ACT NOW CE study

	Error rate (%)	95% PI	*p*-value	I^2
MRA Literature	6.57	0.54–18.5	< 0.0001	0.998
		95% CI		
ACT NOW CE Study (All-Fields)	1.04	0.77–1.34	< 0.0001	0.889
ACT NOW CE Study (Populated-Fields)	2.57	1.88–3.35	< 0.0001	0.897

Note. MRA = Medical Record Abstraction; PI = Prediction Interval; CI = Confidence Interval; I^2 = Higgins and Thompson's I^2 statistic for measuring the degree of heterogeneity, where ≤ 25%, indicating low heterogeneity; 25 – 75% indicating moderate heterogeneity; and > 75%, indicating considerable heterogeneity [37].

Garza *et al. BMC Medical Research Methodology*   (2024) 24:304

Page 5 of 8

these results, it appears that the MRA-QC framework implemented as part of the ACT NOW CE Study was successful in controlling MRA error rates.

Reports of clinical studies in the recent literature routinely lack descriptions of how the quality of the MRA was measured and controlled as well as the error rate ultimately obtained [12]. For clinical studies that did report an error rate, substantial variability was noted in the way error rates were measured, calculated, and expressed [12, 38]. The ACT NOW CE Study was unique in that it implemented and evaluated formalized MRA training and continuous QC processes in an effort to improve data quality [13, 32]. To our knowledge, this was the first time that an MRA-QC framework, such as that published by Zozus and colleagues [12, 13, 32], was implemented and evaluated throughout the course of an ongoing clinical research study. There is a lack of evidence in the literature to suggest that previous clinical studies had implemented any formalized training or QC process to address error rates. As such, for this comparison, we made the decision (1) to limit to comparing against an overall error rate for each study rather than comparing rates across sites or over time; and (2) to provide both conservative (populated-field) and optimistic (all-field) measurements to account for variability across the literature. Given the variability and potential magnitude of the error rates from MRA, researchers should implement a formal data quality control framework that includes prospective quality assurance, such as abstraction guidelines [12] with real-time error checks [40], abstraction training [32], and quality control during the abstraction process. These recommendations are echoed in the GCDMP chapter on *Form Completion Guidelines* [41].

Addressing abstractor-related variability

The reliance on human performance and associated underlying cognitive processes could be responsible for some or all of the variability and could be affected by the level of complexity of the data abstracted for a particular study. For example, the more cumbersome it is to identify, interpret, and collect a specific value from the EHR, the more likely for human error. The amount of abstractor-related variability in abstraction and quality control processes are likely residual effects of the traditional, bespoke, and manual data management techniques that existed within the clinical research and clinical data management professions prior to the last two decades [12, 42, 43]. Fundamentally, we recommend increasing standardization and QC of processes for capturing and processing data by qualified and trained research team members. The SCDM's *Certified Clinical Data Management Exam*™ (CCDM™) assesses a set of universal evidence-based, professional standards for individuals who manage data

from clinical studies [44]. Use of the CCDM™ exam as a tool for establishing competency could reduce variability universally. For example, the CCDM™ exam assesses (in those managing clinical data) the application of evidence-based practice and use of higher-order cognitive abilities (i.e., evaluation, synthesis, creation) [43, 44], potentially reducing variability in processes and human performance that were identified as potential sources of variability.

Addressing performance improvement-related variability

Empirical studies suggest that there is significant variability in the abstraction and quality control processes used; [45, 46] these different methods, process aids, and quality control activities could be responsible for the amount of variation observed in the error rates obtained from the literature. Several authors have further explored these underlying reasons for the high variability in abstraction [45–51]. Further exploration on the causes of this variability is an important area for future research. In particular, the identification of human performance-related sources of variability with training-related root causes versus those caused by the abstraction tools and processes points to improvement interventions [52–54].

A case for MRA guidelines and continuous QC

Although abstraction guidelines constitute a primary mechanism for preventing abstraction errors, they are not often used in clinical studies, which have traditionally relied on form completion guidelines. Until recent recommendations [41], clinical study form completion guidelines traditionally specified definition and format of fields and instructions for documenting exceptions, such as missing values, but usually stopped short of specifying locations in the patient chart from which to pull information, and acceptable alternative locations applicable across multiple clinical sites when the preferred source did not contain the needed data. The MRA-QC framework implemented as part of the data management activities for the ACT NOW CE Study addressed this limitation by developing standardized MRA training and abstraction guidelines with detailed instructions for locating each data point in the patient chart [13, 32]. Further, to account for the variability in clinical charting across institutions, secondary and tertiary locations within the EHR were also provided as alternatives, should the primary location not contain any relevant data. This approach offered clear and consistent instructions for all sites to follow, ensuring greater consistency and accuracy of the data collected.

The continuous QC process implemented as part of the MRA-QC framework, also offered an avenue for further clarification of the abstraction guidelines and periodic check-ins with each site to confirm consistency in interpretation of those guidelines. For example, the

abstraction guidelines were updated significantly after the training [32]. The guidelines were further updated following the routine quality control (independent re-abstraction) events where the root cause of errors was determined to be ambiguities in the abstraction guidelines [13].

A case for quantifying MRA error rates

It is unfortunate that the tendency to associate clinical research with rigorous and prospective data collection further fuels the perception that abstraction or chart review is not a factor in data accuracy when, in fact, (1) the chart itself and manual abstraction from the chart are the sources of most clinical research data error, and (2) manual abstraction from the chart (MRA) remains the most commonly used method for data collection [12, 26, 49, 55–57]. Despite recommendations for measuring and monitoring MRA data quality [46, 50, 58], abstraction error usually remains unquantified in even the most rigorous clinical studies [38, 48, 58, 59]. Based on the now considerable evidence, we echo recommendations in the MRA Framework for abstractor training, tools, conducive environment and ongoing measurement and control of the MRA error rate [12] and add to the calls for reporting data accuracy measures with research results [60]. Reporting a data accuracy measure with research results should be expected in the same way that confidence intervals are expected; it is difficult, and not recommended, to interpret results in their absence.

Limitations

Limitations specific to the analysis of error rates for the ACT NOW CE Study [13] and the comprehensive literature review and meta-analysis [11] have been presented separately and are not repeated here. Briefly, we acknowledge that the error rates for the ACT NOW CE Study were limited to a single, pediatric case study (albeit comprised of 30 individual sites), which may impact the generalizability of those results [13]. As to the comprehensive literature review and meta-analysis, limitations include the following: (1) there is the possibility that we may have missed relevant manuscripts using our search criteria; (2) the results are derived from data that were collected for other purposes (secondary analysis), which could contribute to some of the variability in the error rates; and (3) the results may have slightly less applicability to industry-sponsored studies, as most of the manuscripts identified were from academic organizations and government or foundation-funded endeavors [11]. Still, we do recommend that readers refer to the limitations described in more detail in our previous publications [11, 13] when considering the results presented and the conclusions made in this manuscript.

Similar to the comprehensive literature review, we acknowledge limitations with the identification of MRA-centric manuscripts. As with any literature review, it is possible we may have missed relevant manuscripts due to a lack of standard terminology for data processing methods. Also, because our work is a secondary analysis, it relies on data that were collected for other purposes. Although we used error and field counts reported in the literature, prior work has shown that even these have significant variability [33, 39]. Lastly, we acknowledge the difference between the publication dates of the studies from the literature in comparison to that of our study. The ACT NOW CE Study was conducted in 2018, while the publication dates for the MRA literature ranged from 1987 to 2008. The search was truncated due to inconsistencies in the literature published post-2008, when the use of EDCs became much more widely used. For example, a 2020 review of the EDC literature identified that authors did not consistently report the processes undertaken for collection and processing, nor did they include the error rate [11]. This exclusion of the EDC literature is a potential limitation,

Future direction

While there is general agreement that the validity of research rests on a foundation of data, data collection and processing are sometimes perceived as a clerical part of clinical research. In between rote data entry and scientific validity, however, lie many unanswered questions about effective methods to ensure data quality, which, if answered, will help investigators and research teams balance cost, time, and quality while demonstrating that data are capable of supporting the conclusions drawn.

Conclusion

Based on the comparison of the MRA error rate achieved under formalized quality assurance and process control to those reported in the literature, we conclude that such methods are associated with significantly lower error rates and that measurement and control of the data error rate is possible within a clinical study. We believe that the deployment of our MRA-QC framework allowed the ACT NOW CE Study to maintain error rates significantly lower than the overall MRA error rates identified in the relevant literature.

Abbreviations

MRA	Medical Record Abstraction
QC	Quality Control
ACT NOW CE	Advancing Clinical Trials for Infants with Neonatal Opioid Withdrawal Syndrome Current Experience
NIH	National Institutes of Health
ECHO	Environmental influences on Child Health Outcomes
IDeA	Institutional Development Awards Program
ISPCTN	IDeA States Pediatric Clinical Trials Network
NICHD	National Institute of Child Health and Human Development
NRN	Neonatal Research Network

SCDM	Society for Clinical Data Management
GCDMP	Good Clinical Data Management Practices
CI	Confidence Interval
CCDM™	Society of Clinical Data Management's Certified Clinical Data Management Exam™
EHR	Electronic Health Record
NAACCR	North American Association of Central Cancer Registries
HEDIS	Healthcare Effectiveness Data and Information Set
SDV	Source Data Verification

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12874-024-02424-x.

Supplementary Material 1

Data availability

The datasets generated and/or analyzed during the current study are available in the NICHD Data and Specimen Hub (DASH) repository, https://dash.nichd.nih.gov/study/229026.

Declarations

Ethics approval and consent to participate

The ACT NOW CE Study (parent study; IRB#217689) was reviewed and approved by the University of Arkansas for Medical Sciences (UAMS) Institutional Review Board (IRB). The parent study received a HIPAA waiver, as well as a waiver for informed consent/assent. The quality control study and meta-analysis received a determination of not human subjects research as defined in 45 CFR 46.102 by the UAMS IRB (IRB#239826). All methods were carried out in accordance with relevant guidelines and regulations.

Consent for publication

Not Applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA
[2]University of Texas Health Science Center at San Antonio, Joe R. & Teresa Lozano Long School of Medicine, San Antonio, TX, USA
[3]Department of Biostatistics, University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA
[4]Department of Pediatrics, University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA
[5]University of Colorado Denver, Anschutz Medical Campus, Denver, CO, USA
[6]National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, MD, USA
[7]Department of Pediatrics, School of Medicine, University of Louisville, Louisville, KY, USA
[8]Department of Pediatrics, The Larner College of Medicine, University of Vermont, Burlington, VT, USA
[9]Department of Population Health Sciences, University of Texas Health Science Center at San Antonio, 7703 Floyd Curl Drive, San Antonio, TX 78229, USA

References

1. Forrest WH, Bellville JW. The use of computers in clinical trials. Br J Anaesth. 1967;39:311–9.
2. Helms R. Data quality issues in electronic data capture. Drug Inf J. 2001;35(3):827–37. https://doi.org/10.1177/009286150103500320.
3. Helms RW. A distributed flat file strategy for managing research data. In: Proceedings of the ACM 1980 Annual Conference. ACM '80. Association for Computing Machinery; 1980:279–285. https://doi.org/10.1145/800176.809982
4. Collen MF. Clinical research databases–a historical review. J Med Syst. 1990;14(6):323–44. https://doi.org/10.1007/BF00996713.
5. Knatterud GL, Rockhold FW, George SL, et al. Guidelines for quality assurance in multicenter trials: a position paper. Control Clin Trials. 1998;19(5):477–93. https://doi.org/10.1016/s0197-2456(98)00033-6.
6. Division of Health Sciences Policy, Institute of Medicine (IOM). In: Davis JR, Nolan VP, Woodcock J, Estabrook RW, editors. Assuring data quality and validity in clinical trials for regulatory decision making: workshop report. National Academies; 1999. https://doi.org/10.17226/9623.
7. Redman TC. Data quality for the information age. Artech House; 1996.
8. Redman TC. Data quality: the field guide. Digital; 2001.
9. Batini C, Scannapieco M. Data quality: concepts, methodologies and techniques. Springer; 2006.
10. Lee YW, Pipino LL, Funk JD, Wang RY. Journey to data quality. MIT Press; 2006.
11. Garza MY, Williams T, Ounpraseuth S, et al. Error rates of data processing methods in clinical research: a systematic review and meta-analysis. Published Online Dec. 2022;16. https://doi.org/10.21203/rs.3.rs-2386986/v1.
12. Zozus MN, Pieper C, Johnson CM, et al. Factors affecting accuracy of data abstracted from medical records. PLoS ONE. 2015;10(10):e0138649. https://doi.org/10.1371/journal.pone.0138649.
13. Garza MY, Williams T, Myneni S, et al. Measuring and controlling medical record abstraction (MRA) error rates in an observational study. BMC Med Res Methodol. 2022;22(1):227. https://doi.org/10.1186/s12874-022-01705-7.
14. Young LW, Hu Z, Annett RD, et al. Site-level variation in the characteristics and care of infants with neonatal opioid withdrawal. Pediatrics. 2021;147(1):e2020008839. https://doi.org/10.1542/peds.2020-008839.
15. Kim B. Understanding diagnostic plots for linear regression analysis. University of Virginia Library Research Data Services + Sciences. September 21, 2015. https://data.library.virginia.edu/diagnostic-plots/
16. Vehtari A, Gelman A, Gabry J. Practical bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat Comput. 2017;27(5):1413–32. https://doi.org/10.1007/s11222-016-9696-4.
17. Leave-One-Out Meta-Analysis. http://www.cebm.brown.edu/openmeta/doc/leave_one_out_analysis.html
18. Lovison G, Bellini P. Study on the accuracy of official recording of nosological codes in an Italian regional hospital registry. Methods Inf Med. 1989;28(3):142–7.
19. McGovern PG, Pankow JS, Burke GL, et al. Trends in survival of hospitalized stroke patients between 1970 and 1985. The Minnesota Heart Survey. Stroke. 1993;24(11):1640–8. https://doi.org/10.1161/01.STR.24.11.1640.
20. Steward WP, Vantongelen K, Verweij J, Thomas D, Van Oosterom AT. Chemotherapy administration and data collection in an EORTC collaborative

group–can we trust the results? Eur J Cancer Oxf Engl 1990. 1993;29A(7):943–7. https://doi.org/10.1016/s0959-8049(05)80199-6.

21. Cousley RR, Roberts-Harry D. An audit of the Yorkshire Regional Cleft database. J Orthod. 2000;27(4):319–22. https://doi.org/10.1093/ortho/27.4.319.

22. Moro ML, Morsillo F. Can hospital discharge diagnoses be used for surveillance of surgical-site infections? J Hosp Infect. 2004;56(3):239–41. https://doi.org/10.1016/j.jhin.2003.12.022.

23. Fleiss JL. The statistical basis of meta-analysis. Stat Methods Med Res. 1993;2(2):121–45. https://doi.org/10.1177/096228029300200202.

24. Marín-Martínez F, Sánchez-Meca J. Weighting by inverse variance or by sample size in random-effects meta-analysis. Educ Psychol Meas. 2010;70(1):56–73. https://doi.org/10.1177/0013164409344534.

25. Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softw. 2010;36:1–48. https://doi.org/10.18637/jss.v036.i03.

26. Zozus MN, Kahn MG, Wieskopf N. Data quality in clinical research. In: Richesson RL, Andrews JE, editors. Clinical Research Informatics. Springer; 2019.

27. ACT NOW Current Experience Protocol. (2018). Advancing clinical trials in neonatal opioid withdrawal syndrom (ACT NOW) current experience: infant exposure and treatment, V4.0. Published online 2018.

28. NIH ECHO Program. IDeA states pediatric clinical trials network. NIH ECHO Program. https://www.echochildren.org/idea-states-pediatric-clinical-trials-network/

29. NIH ECHO Program. IDeA States Pediatric Clinical Trials Network Clinical Sites FOA. NIH ECHO Program. 2015. https://www.nih.gov/echo/idea-states-pediatric-clinical-trials-network-clinical-sites-foa

30. NIH. Institutional development award. NIH Division for Research Capacity Building. 2019. https://www.nigms.nih.gov/Research/DRCB/IDeA/Pages/default.aspx

31. NIH. NICHD Neonatal Research Network (NRN). Euinice Kennedy Shriver National Institute of Child Health and Human Development. 2019. https://neonatal.rti.org

32. Zozus MN, Young LW, Simon AE, et al. Training as an intervention to decrease medical record abstraction errors Multicenter studies. Stud Health Technol Inf. 2019;257:526–39.

33. SCDM. Good Clinical Data Management Practices (GCDMP). Published online 2013. https://scdm.org/wp-content/uploads/2019/10/21117-Full-GCDMP-Oct-2013.pdf

34. Basagaña X, Andersen AM, Barrera-Gómez J, et al. Analysis of multicentre epidemiological studies: contrasting fixed or random effects modelling and meta-analysis. Int J Epidemiol. 2018;47. https://doi.org/10.1093/ije/dyy117.

35. Kutner MH, Nachtsheim CJ, Neter J, Li W. Applied linear statistical models. 5th Edition. McGraw-Hill Education; 2004.

36. Montgomery DC, Peck EA, Vining GG. Introduction to linear regression analysis. 6th Edition. Wiley; 2021. https://www.wiley.com/en-jp/Introduction+to+Linear+Regression+Analysis%2C+6th+Edition-p-9781119578727

37. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557–60. https://doi.org/10.1136/bmj.327.7414.557.

38. Rostami R, Nahm M, Pieper CF. What can we learn from a decade of database audits? The Duke Clinical Research Institute experience, 1997–2006. Clin Trials Lond Engl. 2009;6(2):141–50. https://doi.org/10.1177/1740774509102590.

39. Nahm M, Dziem G, Fendt K, Freeman L, Masi J, Ponce Z. Data quality survey results. Data basics. 2004;10:13–9.

40. Eade D, Pestronk M, Russo R, et al. Electronic data capture-study implementation and start-up. J Soc Clin Data Manag. 2021;1(1). https://doi.org/10.47912/jscdm.30.

41. Hills K, Bartlett, Leconte I, Zozus MN. CRF completion guidelines. J Soc Clin Data Manag. 2021;1(1). https://doi.org/10.47912/jscdm.117.

42. McBride R, Singer SW. Introduction [to the 1995 Clinical Data Management Special issue of controlled clinical trials]. Control Clin Trials. 1995;16:S1–3.

43. Zozus MN, Lazarov A, Smith LR, et al. Analysis of professional competencies for the clinical research data management profession: implications for training and professional certification. J Am Med Inf Assoc JAMIA. 2017;24(4):737–45. https://doi.org/10.1093/jamia/ocw179.

44. Williams TB, Schmidtke C, Roessger K, Dieffenderfer V, Garza M, Zozus M. Informing training needs for the revised certified clinical data manager (CCDMTM) exam: analysis of results from the previous exam. JAMIA Open. 2022;5(1):ooac010. https://doi.org/10.1093/jamiaopen/ooac010.

45. Gilbert EH, Lowenstein SR, Koziol-McLain J, Barta DC, Steiner J. Chart reviews in emergency medicine research: where are the methods? Ann Emerg Med. 1996;27(3):305–8. https://doi.org/10.1016/s0196-0644(96)70264-0.

46. Wu L, Ashton CM. Chart review. A need for reappraisal. Eval Health Prof. 1997;20(2):146–63. https://doi.org/10.1177/016327879702000203.

47. Reisch LM, Fosse JS, Beverly K, et al. Training, quality assurance, and assessment of medical record abstraction in a multisite study. Am J Epidemiol. 2003;157(6):546–51. https://doi.org/10.1093/aje/kwg016.

48. Pan L, Fergusson D, Schweitzer I, Hebert PC. Ensuring high accuracy of data abstracted from patient charts: the use of a standardized medical record as a training tool. J Clin Epidemiol. 2005;58(9):918–23. https://doi.org/10.1016/j.jclinepi.2005.02.004.

49. Jansen ACM, van Aalst-Cohen ES, Hutten BA, Büller HR, Kastelein JJP, Prins MH. Guidelines were developed for data collection from medical records for use in retrospective analyses. J Clin Epidemiol. 2005;58(3):269–74. https://doi.org/10.1016/j.jclinepi.2004.07.006.

50. Allison JJ, Wall TC, Spettell CM, et al. The art and science of chart review. Jt Comm J Qual Improv. 2000;26(3):115–36. https://doi.org/10.1016/s1070-3241(00)26009-4.

51. Simmons B, Bennett F, Nelson A, Luther SL. Data abstraction: designing the tools, recruiting and training the data abstractors. SCI Nurs Publ Am Assoc Spinal Cord Inj Nurses. 2002;19(1):22–4.

52. Kirwan B. A guide to practical human reliability assessment. 1st Edition. CRC Press; 1994.

53. Juran JM, Godfrey AB. Juran's quality handbook. 5th Edition. McGraw Hill Professional; 1999.

54. Deming EW. Out of the crisis. 1st MIT Press Edition. MIT Press; 2000.

55. Lee K, Weiskopf N, Pathak J. A Framework for data quality assessment in clinical research datasets. AMIA Annu Symp Proc. 2018;2017:1080–9.

56. Zozus MN, Hammond WE, Green BB, et al. Assessing data quality for healthcare systems data used in clinical research (Version 1.0): an NIH health systems research collaboratory phenotypes, data standards, and data quality core white paper. NIH Collaboratory; 2014. https://dcricollab.dcri.duke.edu/sites/NIHKR/KR/Assessing-data-quality_V1%200.pdf.

57. Blumenstein BA. Verifying keyed medical research data. Stat Med. 1993;12(17):1535–42. https://doi.org/10.1002/sim.4780121702.

58. Nahm ML, Pieper C, Cunningham M. Quantifying data quality for clinical trials using electronic data capture. PLoS ONE. 2008;3:e3049. https://doi.org/10.1371/journal.pone.0003049.

59. Schuyl ML, Engel T. A review of the source document verification process in clinical trials. Drug Inf J. 1999;33:737–84.

60. 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. 2016;4(1). https://doi.org/10.13063/2327-9214.1244

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.