ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 12| Issue: 5| May 2026|| Journal DOI: 10.36713/epra2013 ||SJIF Impact Factor 2026: 8.753 || ISI Value: 1.188

A REVIEW OF METHODS AND STANDARDS IN ENSURING RELIABILITY OF CLINICAL DATA AND CLINICAL DATA ABSTRACTION AND THEIR IMPLICATION IN U.S HEALTHCARE SYSTEM

Omonye Jones Silas[1], Solomon Doe Adjaottor[2]
[1]*University of the Cumberlands, USA*
[2]*Department of Accounting and Finance, Kwame Nkrumah University of Science and Technology, Ghana*

Article DOI: https://doi.org/10.36713/epra27861
DOI No: 10.36713/epra27861

ABSTRACT
There is currently a lack of consistency and generalizability in the methods used to assess electronic health record (EHR) data quality as well as abstracted clinical or EHR data. Using the Boolean function, keywords were used to search and source for articles which were later screened based on their titles and objectives to select only relevant literature materials for this study. A total of nineteen (19) articles were finally obtained and used for this review. The findings from this review show that ensuring reliability in clinical data and clinical data abstraction requires a combination of strong conceptual frameworks, standardized technical processes, and structured human practices. For clinical data, reliability is strengthened through multi-dimensional quality frameworks, standardized pre-processing pipelines, and rigorous validation methods that address missingness, inconsistencies, and temporal complexity. For clinical data abstraction, reliability depends on trained personnel, clear abstraction manuals, continuous quality assurance, and the strategic use of automation supported by human oversight. Manual abstraction remains essential for complex or subjective information, while automated and hybrid approaches improve efficiency for structured data.

KEYWORDS: *Reliability, Clinical data, Abstraction, Medical record abstraction, Electronic health record, Electronic medical records*

INTRODUCTION
Clinical data abstraction (CDA) is the process of identifying and capturing key administrative and clinical data elements. The purpose of abstraction includes the collection of data related to administrative coding functions, quality improvement, patient registry functions, and clinical research (Watzlaf, 2021). Clinical data abstraction, sometimes referred to as medical record abstraction (MRA), has traditionally been, and continues to be, one of the most common forms of data acquisition for clinical research or studies (Kellar *et al*., 2016). The overall quality of healthcare and patient treatment depends heavily on the quality of data. Therefore, having inaccurate, incomplete, and inconsistent data and documentation can result in errors and adverse events (Maurette & Comite, 2002) that may affect patient safety (Abramson *et al*., 2011), limit health information exchange (HIE), and hinder clinical research (Clark *et al*., 2013).

The quality of CDA has often been questioned (Nahm, 2012a; Zozus *et al*, 2015), as it is susceptible to human error (Zozus *et al*., 2014; Zozus *et al*, 2015) and often adds to the overall complexity of clinical research (Sung *et al*., 2003; Getz & Campo, 2018). Clinical data abstraction process remains largely uncharacterized and produces inconsistent results (Nahm, 2012b). Furthermore, CDA is a primary method for validating phenotypes for electronically extracting data from healthcare information systems (Newton *et al*., 2013). Such validation remains subject to the same CDA quality challenges (Zozus *et al*., 2015). It is worth noting that CDA errors are less likely to be detected by downstream data processing, such as data entry or programmatic data cleaning (Zozus *et al*., 2019). For example, an incorrect but plausible value chosen from the medical record will not be detected by valid range checks. Clinical data abstraction errors that result in plausible values will only be detected through comparison to the medical record (i.e., re-abstraction). Accordingly, it is critical that the quality of the data collected through CDA be closely monitored and managed (Zozus *et al*., 2019).

Some studies have shown that error rates associated with CDA are an order of magnitude greater than other data collection methods (Nahm, 2012a; Zozus *et al*, 2015), and the most significant source of error in clinical research (Blumenstein, 1993; Jansen *et al*., 2005). Moreover, the inherent complexities of electronic health records (EHR) often add to the variability of the data abstracted, further contributing to the high and highly variable discrepancy rates associated with CDA (Nahm *et al*., 2008; Zozus *et al*, 2015; Zozus *et al*., 2019; Utomi et al., 2024).
There is currently inconsistency in the methods used to assess EHR data quality as well as abstracted clinical data. In order that the reuse of EHR data for clinical research will be accepted,

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 12| Issue: 5| May 2026|| Journal DOI: 10.36713/epra2013 ||SJIF Impact Factor 2026: 8.753 || ISI Value: 1.188

researchers ought to adopt validated, systematic methods of quality assessment of abstracted clinical data (Weiskopf & Weng, 2013). Therefore, this review aims to highlight and provide commentary on the methods used in ensuring reliability in clinical data and clinical data abstraction.

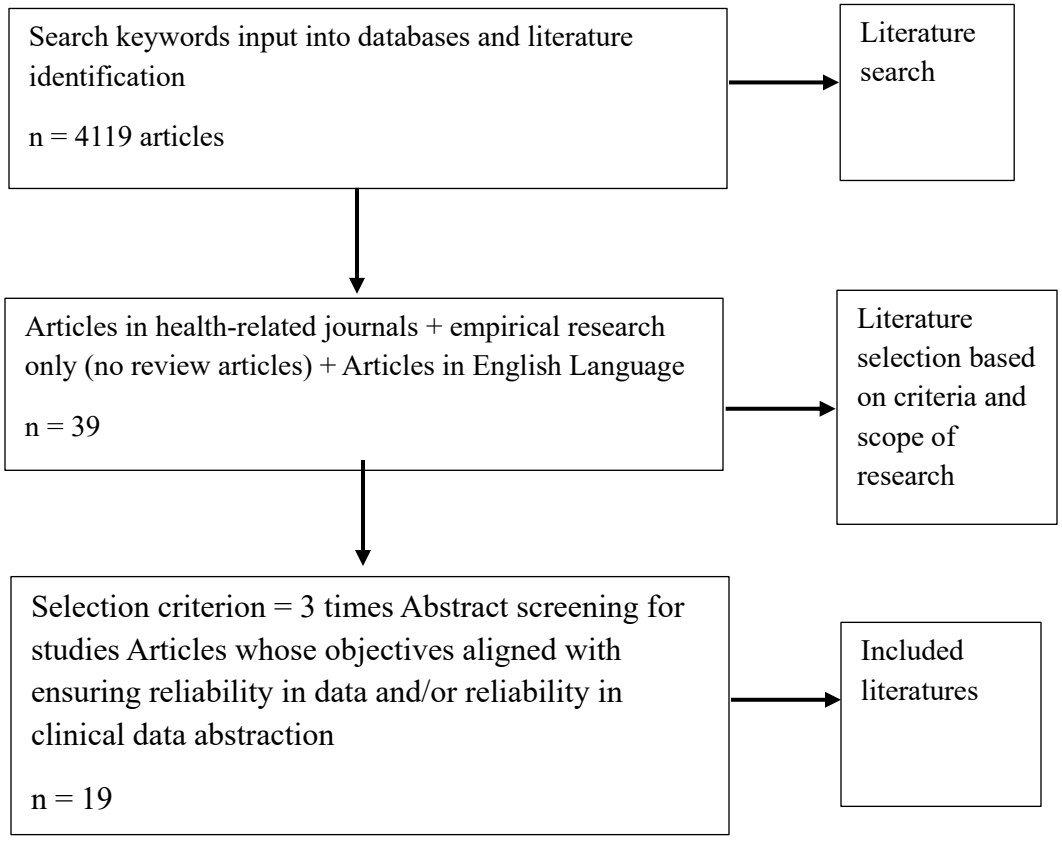
METHODOLOGY

This literature review aims to highlight the methods adopted in ensuring reliability in clinical data abstraction. Drawing exclusively on 19 high-quality empirical and review studies, this review integrates findings across medical data related studies. A systematic literature review approach was adopted in this study. A review protocol was outlined prior to conducting the systematic review, which is in accordance with the review guidelines proposed by Kitchenham *et al*. (2010). The research questions were first defined, after which the keywords identified in this research were input into databases to select the relevant literature materials. The databases that were used to source literature materials in this study included: Science Direct, Google Scholar, and Research Gate. Search keywords or statements used are as follows:

1. Reliability in the methods and standards of clinical data abstraction
2. Methods and strategy of Clinical data abstraction
3. Clinical data abstraction AND Methods OR strategy AND Standards AND Reliability

A selection criteria consisting of documents within health data or clinical data or medical data as well as articles written in English were used to filter the documents and select relevant literature. A total of 39 literature materials were obtained using the above keywords and criteria for searching literature in the various search engines. The 39 literature materials were subjected to three times repeated manual screening based on the objectives of the study. Articles whose objectives aligned with ensuring reliability in data and/or reliability in clinical data abstraction were selected and further advanced in this review study. Additionally, review articles were excluded from the initial 39 literature materials. Out of this number, 19 literature materials were selected after excluding articles based on these criteria.

PRISMA Flow Diagram Used In this Study



*n = number of articles

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 12| Issue: 5| May 2026|| Journal DOI: 10.36713/epra2013 ||SJIF Impact Factor 2026: 8.753 || ISI Value: 1.188

FINDINGS

Hybrid and NLP-Assisted Abstraction Approaches

Hybrid abstraction models that combine manual review with natural language processing (NLP) are emerging as promising approaches to improve efficiency while maintaining high reliability. Watzlaf *et al*. (2021) noted that NLP is increasingly used to process unstructured clinical text, but its reliability depends on algorithm sophistication and the quality of underlying documentation. They emphasize that NLP tools often struggle with ambiguous language, contradictory entries, and context-dependent clinical concepts, making human oversight essential.

Alzu'bi *et al*. (2021) similarly highlight that while NLP can accelerate data extraction, it cannot fully replace human judgment, particularly for complex or nuanced variables. They argue that hybrid models, where NLP performs initial extraction and human abstractors validate and refine outputs, offer the best balance between efficiency and accuracy. These approaches reduce workload while preserving the interpretive strengths of manual abstraction. They further emphasize that effective abstraction requires a combination of clinical knowledge, technical skills, and familiarity with EHR systems. Watzlaf *et al*. (2021) underscore the importance of formal training, credentialing, and ongoing professional development.

Methods and Standards of Ensuring Reliability in Clinical Data

Data quality dimensions and conceptual frameworks

Ensuring the reliability of clinical data begins with a clear understanding of the fundamental dimensions that define data quality. Weiskopf & Weng (2013) provide one of the most influential frameworks, thus, identifying five essential dimensions which include completeness, correctness, concordance, plausibility, and currency. These dimensions capture (i) whether expected data are present, (ii) whether they accurately reflect clinical reality, (iii) whether they agree across sources, (iv) whether they make physiological sense, and (v) whether they are recorded in a timely manner. The review by Weiskopf and Weng (2013) emphasizes that data quality is not absolute but rather "fit for use," meaning that the adequacy of data depends on the specific research or clinical purpose. This perspective underscores the need for transparent reporting of data quality assessment methods to support reproducibility.

Feder (2018) expands on these concepts by articulating five similar domains which include accuracy, completeness, consistency, credibility, and timeliness, and linking them to practical appraisal methods. Feder (2018) highlights the importance of logic-based validation rules, inter-rater reliability testing, and comparison with gold-standard manual chart review, all of which are standards or criteria for assessing the reliability of clinical data. Feder (2018) also stresses the need to evaluate missingness patterns and apply appropriate statistical techniques to mitigate bias. Together, these frameworks establish a conceptual foundation for assessing the reliability of clinical data, emphasizing multi-dimensional evaluation, methodological transparency, and the alignment of data quality assessments with intended use.

Standardized pre-processing pipelines for electronic medical/health record data

Standardized pre-processing pipelines play a critical role in ensuring the reliability of electronic medical record (EMR) or electronic health record (EHR) data, particularly as healthcare systems increasingly rely on machine learning and predictive analytics. Luo *et al*. (2025) introduce Electronic Medical Record Longitudinal Irregular Data Preprocessing (EMR-LIP), a lightweight, modular framework designed to standardize the pre-processing of irregular longitudinal EMR data. EMR-LIP addresses common challenges such as inconsistent sampling intervals, asynchronous measurements, and heterogeneous variable types. By categorizing variables across four dimensions (i.e. time attributes, value types, acquisition methods, and aggregation methods), the framework produces cleaner and more consistent datasets. According to Luo *et al*. (2025), models trained on EMR-LIP-processed data demonstrated superior performance and fairness, with stable results across grouping factors or variables such as ethnic groups and resampling intervals. This highlights the framework's contribution to reproducibility and equity in clinical modeling.

Similarly, Ramakrishnaiah *et al*. (2023) present electronic health record quality control (EHR-QC), an automated pipeline that integrates data standardization with comprehensive quality control. Electronic health record quality control maps raw clinical data to the Observational Medical Outcomes Partnership-Common Data Models and applies systematic procedures to detect and correct missingness, outliers, physiologic implausibility, and unit inconsistencies. Ramakrishnaiah *et al*. (2023) indicated that the pipeline significantly improves predictive model performance and reduces manual pre-processing time, demonstrating scalability across large datasets. These quality control steps substantially improved the performance of predictive models for outcomes such as sepsis and mortality, demonstrating the importance of structured anomaly management in large-scale clinical datasets. Together, EMR-LIP and EHR-QC illustrate how standardized pre-processing pipelines enhance data reliability by reducing variability, improving data integrity, and enabling consistent downstream analytics.

Data validation and quality assessment methods

Reliable clinical data require rigorous validation and systematic quality assessment. Weiskopf and Weng (2013) outline seven (7) methodological approaches for evaluating data quality, including comparison with gold standards, data element agreement, data source agreement, distributional analysis, validity checks, log review, and element presence. These methods allow researchers to detect inconsistencies, implausible values, and structural errors within EHR datasets. Feder (2018) complements this by emphasizing logic-based validation rules, inter-rater reliability testing, and triangulation across multiple data sources. Feder

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 12| Issue: 5| May 2026|| Journal DOI: 10.36713/epra2013 ||SJIF Impact Factor 2026: 8.753 || ISI Value: 1.188

(2018) reinforces the need for rigorous handling of missing and anomalous data, emphasizing that missingness is rarely random and must be evaluated for patterns such as data missing completely at random (MCAR), data missing at random (MAR), or data missing not at random (MNAR). Additionally, Feder (2018) indicated that applying appropriate statistical techniques such as multiple imputation effectively manages anomaly and enhances the accuracy of clinical insights, supports reproducible research, and strengthens the integrity of healthcare analytics.

Liu *et al*. (2021) extend these principles through the San Diego Approach to Variable Validation (SDAVV), a structured framework for validating variables derived from large-scale EHR data. The San Diego Approach to Variable Validation introduces statistically grounded sampling strategies, minimum sample size calculations, and iterative validation cycles using positive predictive value and negative predictive value as primary performance metrics. The framework also provides formulas for estimating sensitivity and specificity when true case status is unknown, a common challenge in large EHR datasets. Empirical applications of SDAVV demonstrated high accuracy in phenotyping tasks, reinforcing the value of structured, prospective validation. Collectively, these studies underscore that robust data validation requires a combination of conceptual rigor, statistical methodology, and iterative refinement to ensure reliability in clinical research and healthcare decision-making.

Subjecting data to abstraction

Automated data abstraction metrics

Automated data extraction has become increasingly important for improving efficiency and scalability in clinical data management, but its reliability depends heavily on standardization and validation. Comparisons between automated and manual abstraction reveal important insights into reliability. Brazeal *et al*. (2021) show that automated abstraction performs well for structured demographic variables, achieving "very good to nearly perfect agreement", but performs poorly for clinical histories and chronic disease variables, where agreement is often only fair or moderate. Automated systems also struggled with missing data, with discrepancies reaching up to 76% for some variables. Altman *et al*. (2018) similarly reported that automated extraction frequently fails to capture nuanced or contradictory information in unstructured clinical notes, leading to systematic under-ascertainment of conditions and potential bias in epidemiological analyses. These discrepancies can introduce bias into epidemiological analyses and reduce statistical power.

Kapoor *et al*. (2021) demonstrated that automated abstraction can be highly reliable when supported by standardized nomenclature and structured data environments. Kapoor *et al*. (2021) developed a platform known as the HINGE platform which successfully automates the Fetch-Transform-Load process and applies TG-263 nomenclature standards to harmonize anatomical labels, improving consistency across institutions. These findings suggest that automated abstraction is most reliable when data are structured and standardized, while manual abstraction remains superior for complex, narrative, or ambiguous information.

Inter-rater and intra-rater agreement metrics in manual abstraction

Manual clinical data abstraction remains a cornerstone of reliable clinical research, and its credibility depends heavily on inter-rater and intra-rater agreement. Mi *et al*. (2013) demonstrate that trained non-physician research coordinators can achieve high agreement with a clinician gold standard, particularly for objective variables such as surgical laterality and co-morbidities. The findings from Mi *et al*. (2013) show that agreement was generally excellent for straightforward clinical parameters, though reliability declined for complex or subjective variables. Seaman *et al*. (2017) similarly report "almost perfect agreement" for objective intensive care unit (ICU) nursing care indicators, reinforcing that structured variables yield higher reliability than interpretive ones. However, they also highlight variability in abstraction accuracy when definitions are ambiguous or documentation is inconsistent. Injeyan *et al*. (2021) add further nuance, showing that while intra-rater reliability tends to be strong, inter-rater reliability can be "fair to moderate" when abstractors interpret clinical narratives differently. Collectively, these studies underscore that manual abstraction can be highly reliable when supported by clear definitions, structured variables, and trained personnel. Yet they also reveal inherent vulnerabilities, particularly for subjective or poorly documented data, that must be addressed through standardization and ongoing quality monitoring. These findings have important implications for the U.S. healthcare system, where manual abstraction continues to support quality reporting, registry development, and clinical research.

Temporal abstraction and longitudinal data integrity

Temporal abstraction is essential for ensuring the reliability of longitudinal clinical data, particularly when raw time-stamped values must be transformed into clinically meaningful patterns. Pham and Duong (2007) describe temporal abstraction as the process of converting primitive time-stamped data into higher-level summaries such as states, trends, and temporal intervals. This transformation supports clinical reasoning by enabling the interpretation of patient trajectories rather than isolated measurements. Pham and Duong (2007) emphasize that temporal abstraction must account for irregular sampling, varying measurement frequencies, and contextual clinical factors to maintain data integrity.

Reliable temporal abstraction requires consistent rules for defining states (e.g., "hypertensive episode"), identifying trends (e.g., "increasing creatinine"), and determining clinically relevant intervals. Pham and Duong (2007) highlight that poorly designed abstraction rules can distort patient histories, leading to inaccurate conclusions. Their work underscores the importance of standardized temporal reasoning frameworks to ensure that longitudinal data remain faithful to clinical reality.

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 12| Issue: 5| May 2026|| Journal DOI: 10.36713/epra2013 ||SJIF Impact Factor 2026: 8.753 || ISI Value: 1.188

Methods and Standards of Ensuring Reliability in Clinical Data Abstraction

Training, standardization, and use of abstraction manuals

Training and standardization are essential for ensuring reliable clinical data abstraction. Pan *et al*. (2005) demonstrate that standardized medical records used as training tools significantly improve abstraction accuracy by helping abstractors identify and correct errors before data collection begins. Their study shows that "standardized charts are effective for identifying and correcting abstraction errors," highlighting the value of gold-standard training materials. Mi *et al*. (2013) reinforce this by showing that structured training protocols and well-defined abstraction manuals enable non-clinicians to achieve high reliability comparable to clinicians. Seaman *et al*. (2017) similarly emphasize that clear definitions and standardized abstraction instruments reduce variability in ICU nursing data abstraction.

Garza *et al*. (2022, 2024) provide strong evidence that structured training combined with systematic quality control dramatically reduces error rates. Garza *et al*. (2024) reports that error rates dropped from 20.5% in literature to just 1.4% when using the medical record abstraction quality control (MRA-QC) framework, hence, demonstrating the transformative impact of standardized training, calibration exercises, and ongoing feedback on clinical data abstraction. Together, these studies show that training is not a one-time event but an ongoing process requiring clear manuals, standardized definitions, and continuous reinforcement. For the U.S. healthcare system where abstraction supports quality metrics, reimbursement, and research, robust training and standardization are indispensable for ensuring reliable and reproducible data.

Quality assurance and quality control in abstraction

Quality assurance (QA) and quality control (QC) are central to ensuring the reliability of clinical data abstraction. Zozus *et al*. (2015) identify 292 factors that influence abstraction accuracy and propose a systems-based framework emphasizing feedback loops, re-abstraction, and continuous monitoring. Their work highlights that abstraction is vulnerable to human, environmental, and methodological factors, making QA/QC essential rather than optional. Seaman *et al*. (2017) similarly emphasize the need for ongoing supervision and periodic reliability checks, noting that even experienced abstractors may drift from standardized definitions over time.

Garza *et al*. (2022) and Garza *et al*. (2024) provide empirical evidence of the effectiveness of structured QA/QC systems. Their medical record abstraction quality control (MRA-QC) framework integrates prospective quality assurance with active monitoring, dramatically reducing abstraction error rates to as low as 1.4%. The framework includes calibration exercises, double-review processes, and systematic error tracking, demonstrating that rigorous QC can significantly enhance reliability. These studies collectively show that QA/QC must be embedded throughout the abstraction lifecycle, from training to data collection to final review of the clinical data. In the U.S. healthcare system, where abstracted data inform regulatory reporting, clinical quality measures, and research, robust QA/QC processes are essential to ensure accuracy, reduce bias, and maintain trust in clinical datasets.

Organizational and workforce standards for abstraction

Organizational structures and workforce competencies play a critical role in ensuring reliable clinical data abstraction. Watzlaf *et al*. (2021) report that healthcare organizations use decentralized abstraction models embedded within coding or quality departments. While this structure supports workflow integration, it also introduces variability in training, oversight, and abstraction practices. The findings by Watzlaf *et al*. (2021) highlight the need for standardized organizational policies, clear role definitions, and consistent competency requirements to ensure reliability across institutions.

DISCUSSION

The reliability of clinical data and clinical data abstraction is increasingly recognized as a foundational requirement for high-quality healthcare delivery, research, and policy decision-making. Across the literatures used, a consistent theme is that reliability is not achieved through a single method but through the integration of conceptual frameworks, standardized processes, technological tools, and human expertise.

With respect to methods and standards for ensuring reliability in clinical data, a key insight is the shift toward multi-dimensional and purpose-driven definitions of data quality. Rather than treating reliability as a binary attribute, contemporary approaches such as those used by Garza *et al*. (2022) and Garza *et al*. (2024) emphasize that data must be evaluated in relation to its intended use. This perspective has important implications for both research and clinical practice, as it requires transparency in how data quality is assessed and reported. It also reinforces the need for alignment between data quality metrics and analytical objectives, particularly in complex environments such as electronic health records (EHRs), where variability and incompleteness are inherent.

The growing reliance on large-scale EHR data has further highlighted the importance of standardized pre-processing pipelines. These pipelines serve as critical intermediaries between raw clinical data and downstream analytical applications, reducing heterogeneity and enhancing consistency. Their role extends beyond simple data cleaning to include the harmonization of temporal structures, variable definitions, and measurement units. Importantly, the integration of automated quality control mechanisms within these pipelines demonstrates how scalability and reliability can be achieved simultaneously. The evidence suggests that without such standardization, even advanced analytical techniques may yield biased or non-reproducible results. Equally significant is the role of rigorous validation and quality assessment methodologies. Reliable clinical data cannot be assumed but must be systematically verified through multiple complementary approaches. The use of gold standards, statistical validation techniques, and iterative refinement processes reflects a maturing field that increasingly values methodological rigor.

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 12| Issue: 5| May 2026|| Journal DOI: 10.36713/epra2013 ||SJIF Impact Factor 2026: 8.753 || ISI Value: 1.188

Furthermore, the recognition that missing data is often non-random (Feder, 2018) introduces an additional layer of complexity, necessitating sophisticated statistical handling to avoid biased inferences. Together, these approaches highlight that data reliability is an ongoing process requiring continuous evaluation rather than a one-time verification step.

While automated abstraction offers clear advantages in terms of efficiency and scalability, its limitations in handling unstructured and context-dependent information remain evident. This has led to the emergence of hybrid models that leverage the strengths of both machine-driven and manual approaches. Such models acknowledge that while algorithms can process large volumes of data rapidly, human abstractors are indispensable for interpreting ambiguity and ensuring contextual accuracy (Alzu'bi *et al*., 2021; Watzlaf *et al*., 2021). This balance is likely to define the future of clinical data abstraction, particularly as natural language processing technologies continue to evolve. The reliability of abstraction is also closely tied to the competencies of the workforce and the systems within which they operate. Training, standardization, and the use of detailed abstraction manuals emerge as critical enablers of consistency. It is worth noting that training is not a static requirement but a continuous process supported by feedback, calibration, and performance monitoring. This is particularly relevant given the variability introduced by subjective interpretation, especially when dealing with complex clinical narratives. Standardization efforts, therefore, play a crucial role in reducing ambiguity and aligning abstractors' interpretations.

Quality assurance and quality control mechanisms further reinforce the reliability of abstraction processes. The evidence strongly supports the integration of systematic quality assurance or quality control frameworks throughout the abstraction lifecycle. Continuous monitoring, re-abstraction, and error tracking not only identify inaccuracies (Zozus *et al*., 2015) but also prevent their recurrence through feedback loops. This systems-based approach reflects an understanding that errors in abstraction are influenced by a wide range of factors, including human, organizational, and methodological variables.

Finally, organizational and structural considerations are essential for sustaining reliable abstraction practices. Variability in institutional policies, training protocols, and oversight mechanisms can lead to inconsistencies in data quality across settings. This highlights the need for standardized organizational frameworks that define roles, competencies, and performance expectations. In large and complex healthcare systems, such standardization is critical to ensuring that data generated across different units and institutions remain comparable and trustworthy.

Implications for the U.S. healthcare system

The methods and standards identified in this review have significant implications for the U.S. healthcare system, where reliable data underpin quality reporting, reimbursement, clinical decision support, and research. The conceptual frameworks and pre-processing pipelines used to ensure reliability in clinical data directly support national priorities such as interoperability, data standardization, and equitable analytics. Tools like EMR-LIP and EHR-QC help reduce variability across institutions, enabling more consistent performance measurement and improving the reliability of predictive models used in clinical care and population health management.

The findings on clinical data abstraction highlight the importance of investing in workforce training, standardized manuals, and QA/QC systems. As U.S. healthcare organizations rely heavily on abstracted data for accreditation, regulatory reporting, and value-based payment programs, the demonstrated benefits of structured training and rigorous QC such as the dramatic error reduction shown by Garza *et al*. (2024) underscore the need for national standards and institutional accountability, thus reliability in clinical data needs to be prioritized.

In the context of U.S. healthcare, where EMR systems generate vast amounts of time-dependent data, robust temporal abstraction methods are critical for supporting clinical decision support, predictive modeling, and population health analytics. By ensuring that longitudinal data are interpreted consistently and accurately, temporal abstraction strengthens the reliability of downstream analyses and enhances the utility of EMR data for research and care delivery. The comparison between automated and manual abstraction has practical implications for scalability and cost. While automation can reduce workload and support real-time surveillance, its limitations with unstructured data mean that hybrid models are necessary to maintain accuracy. This aligns with broader U.S. healthcare trends toward integrating AI tools with human expertise. However, this cannot be overlooked from the governance aspect as Osifowokan *et al*. (2025a, 2025b) indicated that effective health data governance requires the integration of regulatory frameworks, institutional structures, technological systems, and operational practices.

Overall, the reviewed methods highlight reliable, efficient, and equitable health data ecosystem. They highlight the need for coordinated national strategies that combine technical standardization, workforce development, and continuous quality monitoring to strengthen the integrity of U.S. healthcare data.

CONCLUSION

This review shows that ensuring reliability in clinical data and clinical data abstraction requires a combination of strong conceptual frameworks, standardized technical processes, and structured human practices. Manual abstraction remains essential for complex or subjective information, while automated and hybrid approaches improve efficiency for structured data. Hybrid models, where NLP performs initial extraction and human abstractors validate and refine outputs, offer one of the best balances between efficiency and accuracy. For clinical data, reliability is strengthened through multi-dimensional quality frameworks, reliable data abstraction, standardized pre-processing pipelines, and rigorous validation methods that

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 12| Issue: 5| May 2026|| Journal DOI: 10.36713/epra2013 ||SJIF Impact Factor 2026: 8.753 || ISI Value: 1.188

address missingness, inconsistencies, and temporal complexity. For clinical data abstraction, reliability depends on trained personnel, clear abstraction manuals, abstraction metrics, continuous quality assurance, and the strategic use of automation supported by human oversight.

REFERENCES

1. *Abramson, E. L., et al. (2011). Transitioning between electronic health records: Effects on ambulatory prescribing safety. Journal of General Internal Medicine, 26(8), 868–874. https://doi.org/10.1007/s11606-011-1662-z*
2. *Altman, M. R., Colorafi, K., & Daratha, K. B. (2018). The reliability of electronic health record data used for obstetrical research. Applied Clinical Informatics, 9(1), 156–162. https://doi.org/10.1056/s-0038-1627475*
3. *Alzu'bi, A. A., Watzlaf, V. J. M., & Sheridan, P. (2021). Electronic health record (EHR) abstraction. Perspectives in Health Information Management, Spring 2021, 1–16.*
4. *Blumenstein, B. A. (1993). Verifying keyed medical research data. Statistics in Medicine, 12(17), 1535–1542. https://doi.org/10.1002/sim.4780121702*
5. *Brazeal, J. G., Alekseyenko, A. V., Li, H., Fugal, M., Kirchoff, K., Marsh, C., Lewin, D. N., Wu, J., Obeid, J., & Wallace, K. (2021). Assessing quality and agreement of structured data in automatic versus manual abstraction of the electronic health record for a clinical epidemiology study. Research Methods in Medicine & Health Sciences, 2(4), 168–178. https://doi.org/10.1177/26320843211061287*
6. *Clark, J. S., et al. (2013). Assessing and improving EHR data quality (updated). Journal of AHIMA, 84(3), 48–53.*
7. *Feder, S. L. (2018). Data quality in electronic health records research: Quality domains and assessment methods. Western Journal of Nursing Research, 40(5), 753–766. https://doi.org/10.1177/0193945916689084*
8. *Garza, M. Y., Williams, T. B., Ounpraseuth, S., Hu, Z., Lee, J., Snowden, J., Walden, A. C., Simon, A. E., Devlin, L. A., Young, L. W., & Zozus, M. N. (2024). Comparing medical record abstraction (MRA) error rates in an observational study to pooled rates identified in the data quality literature. BMC Medical Research Methodology, 24(1), 304. https://doi.org/10.1186/s12874-024-02424-x*
9. *Garza, M. Y., Williams, T., Myneni, S., Fenton, S. H., Ounpraseuth, S., Hu, Z., Lee, J., Snowden, J., Zozus, M. N., Walden, A. C., Simon, A. E., McClaskey, B., Sanders, S. G., Beauman, S. S., Ford, S. R., Malloch, L., Wilson, A., Devlin, L. A., & Young, L. W. (2022). Measuring and controlling medical record abstraction (MRA) error rates in an observational study. BMC Medical Research Methodology, 22(1), 227. https://doi.org/10.1186/s12874-022-01705-7*
10. *Getz, K. A., & Campo, R. A. (2018). New benchmarks characterizing growth in protocol design complexity. Therapeutic Innovation & Regulatory Science, 52(1), 22–28. https://doi.org/10.1177/2168479017713039.*
11. *Injeyan, H. S., Hogg-Johnson, S., Abdulla, S., Chow, N., Cox, J., Ridding, A., & Jacobs, C. (2021). Intra- and inter-rater reliability of an electronic health record audit used in a chiropractic teaching clinic system: An observational study. BMC Health Services Research, 21(1), 750. https://doi.org/10.1186/s12913-021-06745-1*
12. *Jansen, A. C., van Aalst-Cohen, E. S., Hutten, B. A., Buller, H. R., Kastelein, J. J. P., & Prins, M. H. (2005). Guidelines were developed for data collection from medical records for use in retrospective analyses. Journal of Clinical Epidemiology, 58(3), 269–274. https://doi.org/10.1016/j.jclinepi.2004.07.006*
13. *Kapoor, R., Sleeman, W. C., IV, Nalluri, J. J., Turner, P., Bose, P., Cherevko, A., Srinivasan, S., Syed, K., Ghosh, P., Hagan, M., & Palta, J. R. (2021). Automated data abstraction for quality surveillance and outcome assessment in radiation oncology. Journal of Applied Clinical Medical Physics, 22(7), 256–267. https://doi.org/10.1002/acm2.13308*
14. *Kellar, E., Bornstein, S. M., Caban, A., Celingant, C., Crouthamel, M., Johnson, C., et al. (2016). Optimizing the use of electronic data sources in clinical trials: The landscape, part 1. Therapeutic Innovation & Regulatory Science, 50(6), 682–696. https://doi.org/10.1177/2168479016670689*
15. *Liu, L., Bustamante, R., Earles, A., Demb, J., Messer, K., & Gupta, S. (2021). A strategy for validation of variables derived from large-scale electronic health record data. Journal of Biomedical Informatics, 121, 103879. https://doi.org/10.1016/j.jbi.2021.103879*
16. *Luo, J., Huang, S., Lan, L., Yang, S., Cao, T., Yin, J., Qiu, J., Yang, X., Guo, Y., & Zhou, X. (2025). EMR-LIP: A lightweight framework for standardizing the preprocessing of longitudinal irregular data in electronic medical records. Computer Methods and Programs in Biomedicine, 259, 108521. https://doi.org/10.1016/j.cmpb.2024.108521*
17. *Maurette, P., & Comité Analyse et Maîtrise du Risque de la SFAR. (2002). Ann Fr Anesth Reanim [Annales Françaises d'Anesthésie et de Réanimation]. Annales Françaises d'Anesthésie et de Réanimation, 21(6), 453–454. https://doi.org/10.1016/S0750-7658(02)00645-X*
18. *Mi, M. Y., Collins, J. E., Lerner, V., Losina, E., & Katz, J. N. (2013). Reliability of medical record abstraction by non-physicians for orthopedic research. BMC Musculoskeletal Disorders, 14(1), 181. https://doi.org/10.1186/1471-2474-14-181*
19. *Nahm, M. (2012a). Data quality in clinical research. In R. L. Richesson & J. E. Andrews (Eds.), Clinical research informatics (pp. 175–202). Springer..*
20. *Nahm, M. (2012b). Measuring data quality. In Good clinical data management practices (GCDMP) (Version 2000–present). Society for Clinical Data Management. Available from: https://www.scdm.org/gcdmp. Accessed Aug 2020*
21. *Nahm, M. L., Pieper, C. F., & Cunningham, M. M. (2008). Quantifying data quality for clinical trials using electronic data capture. PLOS ONE, 3(8), e3049. https://doi.org/10.1371/journal.pone.0003049*
22. *Newton, K. M., Peissig, P. L., Kho, A. N., Bielinski, S. J., Berg, R. L., Choudhary, V., et al. (2013). Validation of electronic medical record-based phenotyping algorithms: Results and lessons learned from the eMERGE network. Journal of the American Medical Informatics Association, 20(e1), e147–e154. https://doi.org/10.1136/amiajnl-2012-000896*
23. *Osifowokan, A. S., Agbadamasi, T. O., Adukpo, T. K., & Mensah, N. (2025). Regulatory and legal challenges of artificial intelligence in the US healthcare system: Liability, compliance, and patient safety. World Journal of Advanced Research and Reviews, 25(3), 949-955.*
24. *Osifowokan, A. S., Ahmed, Z., Adukpo, T. K., & Mensah, N. (2025). Enhancing data compliance in the United States healthcare system:*

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 12| Issue: 5| May 2026|| Journal DOI: 10.36713/epra2013 ||SJIF Impact Factor 2026: 8.753 || ISI Value: 1.188

Addressing challenges in HIPAA and HITECH Act implementation. EPRA International Journal.

25. *Pan, L., Fergusson, D., Schweitzer, I., & Hebert, P. C. (2005). Ensuring high accuracy of data abstracted from patient charts: The use of a standardized medical record as a training tool. Journal of Clinical Epidemiology, 58(9), 918–923. https://doi.org/10.1016/j.jclinepi.2005.02.004*
26. *Pham, V. C., & Duong, T. A. (2007). Applying temporal abstraction in clinical databases. 2007 IEEE International Conference on Research, Innovation and Vision for the Future, 154–161. https://doi.org/10.1109/RIVF.2007.369156*
27. *Ramakrishnaiah, Y., Macesic, N., Webb, G. I., Peleg, A. Y., & Tyagi, S. (2023). EHR-QC: A streamlined pipeline for automated electronic health records standardisation and preprocessing to predict clinical outcomes. Journal of Biomedical Informatics, 147, 104509. https://doi.org/10.1016/j.jbi.2023.104509*
28. *Seaman, J. B., Evans, A. C., Sciulli, A. M., Barnato, A. E., Sereika, S. M., & Happ, M. B. (2017). Abstracting ICU nursing care quality data from the electronic health record. Western Journal of Nursing Research, 39(9), 1271–1288. https://doi.org/10.1177/0193945916665814*
29. *Sung, N. S., Crowley, W. F., Genel, M., Salber, P., Sandy, L., Sherwood, L. M., et al. (2003). Central challenges facing the national clinical research enterprise. JAMA, 289(10), 1278–1287. https://doi.org/10.1001/jama.289.10.1278*
30. *Utomi, E., Osifowokan, A. S., Donkor, A. A., & Yowetu, I. A. (2024). Evaluating the Impact of Data Protection Compliance on AI Development and Deployment in the US Health sector. World Journal of Advanced Research and Reviews, 24(2), 1100-1110.*
31. *Watzlaf, V. J. M., Sheridan, P. T., Alzu'bi, A. A., & Chau, L. (2021). Clinical data abstraction: A research study. Perspectives in Health Information Management, Spring 2021, 1–18.*
32. *Weiskopf, N. G., & Weng, C. (2013). Methods and dimensions of electronic health record data quality assessment: Enabling reuse for clinical research. Journal of the American Medical Informatics Association, 20(1), 144–151. https://doi.org/10.1136/amiajnl-2011-000681*
33. *Zozus, M. N., Hammond, W. E., Green, B. G., et al. (2014). 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 Health Care Systems Research Collaboratory. https://www.nihcollaboratory.org/Products/Assessing-data-quality_V1%200.pdf*
34. *Zozus, M. N., Pieper, C., Johnson, C. M., Johnson, T. R., Franklin, A., Smith, J., & Zhang, J. (2015). Factors affecting accuracy of data abstracted from medical records. PLOS ONE, 10(10), e0138649. https://doi.org/10.1371/journal.pone.0138649*
35. *Zozus, M. N., Young, L. W., Simon, A. E., et al. (2019). Training as an intervention to decrease medical record abstraction errors in multicenter studies. Studies in Health Technology and Informatics, 257, 526–539..*