This is a "preproof" accepted article for *Journal of Clinical and Translational Science*.
This version may be subject to change during the production process.
10.1017/cts.2024.551

Source Data Verification (SDV) Quality in Clinical Research: A Scoping Review

Muayad Hamidi, MD[1]; Eric L. Eisenstein, DBA[2]; Maryam Y. Garza, PhD, MPH, MMCi[3]; Kayla Joan Torres Morales, MA[1]; Erika M. Edwards, PhD, MPH[4]; Mitra Rocca, MSc[3]; Amy Cramer, MMCi[5]; Gurparkash Singh, PhD[5]; Kimberly A. Stephenson-Miles, MSc[6]; Mahanaz Syed, PhD[1]; Zhan Wang, PhD[1] ; Holly Lanham, PhD[1]; Rhonda Facile, MS[9] ; Justine M. Pierson, MS[7] ; Cal Collins[7] ; Henry Wei, MD[8] ; Meredith Zozus, PhD[1]

[1]University of Texas Health Science Center at San Antonio,
[2]Duke University,
[3]University of Arkansas for Medical Sciences,
[4]University of Vermont, Vermont Oxford Network,
[5]Janssen R&D LLC,
[6]ICON PLC employee on an assignment to Janssen R&D LLC,
[7]OpenClinica,
[8]Regeneron,
[9]Clinical Data Interchange Standards Consortium (CDISC)

Corresponding Author: Muayad Hamidi, MD, Joe R. and Teresa Lozano Long School of Medicine, University of Texas Health Science Center San Antonio, 7703 Floyd Curl Drive, San Antonio, TX 78229, 210-567-0841, maallah@uthscsa.edu

Conflicts of interest: The authors report no conflicts of interest.

Abstract

Introduction: The value of Source Data Verification (SDV) has been a common theme in the applied Clinical Translational Science literature. Yet, few published assessments of SDV quality exist even though they are needed to design risk-based and reduced monitoring schemes. This review was conducted to identify reports of SDV quality, with a specific focus on accuracy.

Methods: A scoping review was conducted of the Source Data Verification (SDV) and clinical trial monitoring literature to identify articles addressing SDV quality. Articles were systematically screened and summarized in terms of research design, SDV context, and reported measures.

Results: The review found significant heterogeneity in underlying SDV methods, domains of SDV quality measured, the outcomes assessed, and the levels at which they were reported. This variability precluded comparison or pooling of results across the articles. No absolute measures of SDV accuracy were identified.

Conclusions: A definitive and comprehensive characterization of SDV process accuracy was not found. Reducing the SDV without understanding the risk of critical findings going undetected, i.e., SDV sensitivity, is counter to recommendations in Good Clinical Practice and the principles of quality by design. Reference estimates (or methods to obtain estimates) of SDV accuracy are needed to confidently design risk-based, reduced SDV processes for clinical studies.

Keywords: Source Data Verification, SDV, Quality, Clinical Research, Clinical Trial Monitoring

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

Introduction

Clinical trial complexity continues to rise and increases the effort required at clinical investigational sites.[1-5] This has contributed to clinical trial operational inefficiency being considered one of the major impediments to Clinical and Translational Research (CTR).[13] The rate of clinical trial cost increase – driven by study complexity and operational inefficiency – is greater than inflation for other segments of the economy.[14-16] An estimated 46% percent of on-site monitoring time has been attributed to Source Data Verification (SDV).[17] Estimates for the portion of clinical trial costs attributable to SDV range from 25-40%, implicating SDV as a major cost driver.[17-20] Reducing the amount of manual SDV in clinical trials would create an opportunity to increase operational efficiency and lower clinical trial costs.

Box 1: Key Definitions Relevant to Source Document Verification

Monitoring: The act of overseeing the progress of a clinical study, and ensuring that it is conducted, recorded, and reported in accordance with the protocol, standard operating procedures (SOP), Good Clinical Practice (GCP), and the applicable regulatory requirements.[6] Clinical study monitoring includes activities such as ensuring good communication between site investigators and the Sponsor, verifying adequate resources, storage, and accountability of the investigational product and biological samples, Informed Consent, and regulatory compliance and protocol adherence at sites, ensuring that site personnel is qualified for their roles on the study, tracking recruitment, enrollment, and retention, ensuring appropriate reporting of adverse events, deviations and problems, and ensuring the completeness and accuracy of study data through SDV.[7]

Risk-based Monitoring (RBM): an adaptive approach to clinical trial monitoring that directs monitoring focus and activities to the evolving areas of greatest need that have the most potential to impact subject safety and data quality.[8]

On-site monitoring: An in-person evaluation carried out by sponsor personnel or representatives at the sites at which the clinical investigation is being conducted.[9] In risk-based approaches, on-site monitoring may be focused on activities that are critical to safety or study results; this is referred to as *Targeted On-site Monitoring*.[6]

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

Off-site Monitoring: Remote monitoring via telephone and e-mail without visiting the site institutions.[10]

Centralized monitoring: (1) A remote evaluation carried out by sponsor personnel or representatives at a location other than the sites at which the clinical investigation is being conducted.[9] (2) Document review, data review, and analysis performed remotely from the investigator site to examine the data collected to check compliance and identify unusual patterns and deviations.[11] This is also called *remote monitoring*.

Source Data Verification (SDV): the comparison of study data to their original recording to ensure that, "the reported trial data are accurate, complete, and verifiable from source documents."[6]

Source Document Review (SDR)[8] describes the review of source documents for protocol adherence, quality of documentation, as well as site processes in contrast to transcription checking, referred to as Source Data Verification (SDV).

Targeted SDV: A risk-based approach that focuses SDV efforts toward data that are critical to safety or study results. Targeted SDV may result in fewer data fields being verified and is sometimes called *Reduced SDV*.

Reduced SDV: a decrease in the amount of SDV performed for a study through performing SDV for some (or none) of the sites, patients, visits, data elements, or data values.

Quality by Design (QbD): A systematic approach, "that begins with predefined objectives and emphasizes product and process understanding and process control, based on sound science and quality risk management".[12]

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

The comparison of study data to the medical record (or other sources) to verify that the medical record data are accurately reflected in the study data is called Source Data Verification (SDV) (definitions Box 1). Since the earliest reported use in the year 1746,[21] SDV has been used as a tool to find and fix errors from medical record abstraction (MRA), the process associated with the largest error rate in data processing.[22]

Extensive and often 100%, SDV was historically considered necessary to ensure study data quality and has served as the foundation upon which trialists claimed data accuracy and authenticity.[23,24] Cognitively similar to MRA, SDV is a manual inspection process performed by humans, and the error rate is likely similarly high. Finding the twenty-one differences between Figure 1.a and 1.b illustrates challenges with SDV as a mechanism to identify data errors.

Decades ago, the field of HRA demonstrated that errors in the manual inspection process had a significant and often overlooked impact on the average outgoing quality.[25] SDV has long been used as a tool to assure data quality. However, the quality of this tool itself was rarely examined. Ironically, maintaining a current calibration record for all devices utilized in a clinical trial is required, while no comparable requirement exists for the more fallible processes of manual abstraction from medical record sources and manual inspection by SDV. Though at reduced levels through risk-based approaches, SDV is still being used as a major strategy to identify errors and assure high-quality data.[26-28] This thinking must change. As a fallible manual process, SDV was never capable of assuring high-quality data from a fallible source.[25,29] SDV provided a convenient answer, “all data were verified against the source”, but the assumption that often followed, “therefore the data are correct”, was flawed. This leaves us with the uncomfortable difficulty of needing to verify the accuracy of data in the absence of a true gold standard, a conundrum that likely perpetuated reliance on manual SDV.

Reports of SDV error and omission started to emerge with studies evaluating various risk-based monitoring approaches,[30,31] and along with them, acknowledgment that manual SDV processes are not capable of producing error-free data. In this context, the value of the traditional 100% SDV has increasingly been questioned.[18,19,32-34] In a recent national study, respondents characterized the "amount of money, time and resources spent on certain monitoring methods” such as extensive on-site SDV, as being wasteful for commercial trials.[35] Federally funded and investigator-initiated studies traditionally employed more limited monitoring approaches due to budget constraints.[36-39] With the shift toward risk-based approaches articulated in the United

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

States Food and Drug Administration's (FDA's) guidance on risk-based monitoring in 2013[9] and the 2018 revision of the Good Clinical Practice (GCP) guidelines,[40] industry and academic monitoring practices are converging with risk-based approaches now strongly encouraged by prominent industry groups[8,41,42] and regulators.[9,11,43] Prior to the COVID-19 pandemic, however, industry adoption of risk-based monitoring practices, such as off-site remote monitoring, reduced Source Document Review (SDR), reduced SDV, and centralized monitoring (definitions in Box 1), remained low, 25%, 16%, 20%, and 16% respectively.[44]

Only moderate evidence supports the comparability of traditional (100%) and reduced SDV.[27] European Medicines Agency (EMA) has reached beyond the status quo and called for demonstration that alternate monitoring methods are non-inferior to traditional methods.[45] GCP guidelines and the FDA, through Quality by Design (QbD) principles,[6,41] have communicated general heightened expectations for quality planning and the design or selection of study processes with the capability to deliver the data accuracy required to support planned study analyses. However, computational models for the *a priori* design of capable reduced SDV processes are not common, and the inputs needed for the computational models, such as error rates for SDV processes, have not been widely reported. The resulting uncertainty in process capability is a likely contributor to the slow adoption of reduced SDV, along with a lack of comparative evidence, lingering concerns of feasibility, the effort required to change existing organizational processes, and methodological questions regarding reduced SDV. Filling these knowledge gaps will provide the means to design and optimize SDV for clinical studies. We conducted this review to minimize this knowledge gap.

Objective

The goals of this scoping review were (1) to identify published reports of SDV quality and (2) to characterize and summarize the evidence regarding the quality of the SDV performed in clinical studies.

Methods

The Preferred Reporting Items for Systematic Reviews and Meta-analysis extension for Scoping Reviews (PRISMA-ScR)[46] methodology was adopted (Supplementary Material 1 - PRISMA-ScR-Fillable-Checklist). The protocol components, as outlined in the PRISMA-Scr were standardized *a priori* and used by authors, though the protocol was not registered in a review protocol registry such as Open Science Framework.

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

Eligibility criteria: Studies were included if (1) the main focus was the quality of the SDV process, i.e., problems with or ways to achieve the quality of the SDV process, (2) the full-text article was available, and (3) the article was written in the English language. Studies were excluded if they (1) were not focused on SDV quality, (2) were not full-text peer-reviewed articles, e.g., abstracts, posters, regulations, and policies, or (3) were not in English. All types of study designs were included; however, policy documents, guidelines, and regulations were excluded.

Information sources: The following databases were searched: MEDLINE, CINAHL, PsycINFO, Embase, Web of Science, Scopus, Cochrane, The Association of Clinical Research Professionals (ACRP), The Society of Clinical Research Associates (SOCRA), Applied Clinical Trials (ACT), and Google Scholar was used as a safety net to confirm that our search process didn't miss any relevant papers. The last search was executed in October 2022. In addition, the list of similar articles suggested by the NLM website next to each search result.

Search: The following PubMed query was used: ("SDV"[Title/Abstract] OR "clinical trial monitoring"[Title/Abstract]) AND ("clin res"[Journal] OR "clinical research"[Title/Abstract] OR "clinical trial"[Title/Abstract] OR "clinical study"[Title/Abstract]) with the syntax modified as needed to run on the other databases.

Selection of sources of evidence the database searches were performed by three authors (MM, KTM, MNZ), with the initial screening of titles and abstracts performed by two authors (MM, MNZ). The articles that passed the initial screening were retrieved for full-text review. Two independent people reviewed the full text of the retrieved articles and attempted consensus on disagreements. Disagreements for which consensus could not be achieved were adjudicated by a majority vote of authors on weekly adjudication calls.

Data extraction process: A spreadsheet (Supplementary Material 2 - Information abstracted from articles) was created to capture the data needed to be extracted. All authors independently participated in abstracting a predefined list of relevant Information from each included article.

Data items: information abstracted from the included articles is listed in (Table 1) which can be grouped into: article metadata, information about research design and context, and information about SDV accuracy.

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

Critical appraisal of individual sources of evidence: As a scoping review, we sought to characterize all available evidence regarding SDV quality. Since we anticipated a wide variety of evidence, we critically apprised the strength of evidence according to the research design and context of the study.

Synthesis of results: We grouped the studies by study design and context. Examples include prospective *versus* retrospective approach and experimental *versus* quasi- experimental, pre-experimental or descriptive designs. Use of comparators and control, and whether the study measured SDV discrepancies or SDV errors were also used to categorize studies as well as aspects of the research context such as the number of studies, sites, or participants included in the quality assessment, the therapeutic area in which the SDV quality assessment was conducted, and whether the study was industry-funded.

Results

Selection of sources of evidence: The search of bibliographic databases yielded 683 records, and an additional 72 were identified from searching citations (paper references). (Figure 2). From the total of 755 records identified, there were 207 duplicate records removed, 360 excluded per the Title and Abstract screening process, and 6 unretrievable. 182 articles (110 plus 72 in Figure 2) were retrieved and underwent full-text review. The full-text review eligibility assessment resulted in the exclusion of 154 articles (91 plus 63 in Figure 2) and the inclusion of 28 articles. Sixteen of the included articles reported quantitative results.[7,8,31,33,37,38,47-56] Twelve articles reported non-quantitative studies [26,27] [23,28,32,57-63] (Figure 2).

Characteristics of sources of evidence:

Quantitative Studies

Seven of the included quantitative studies prospectively collected data after assigning the patient, site, or study to some monitoring strategy[7,37,47,48,52,54,55]. Nine analyzed existing data[8,31,33,38,49-51,53,56] (references detailed in Table 2). Ten studies were classified as using pre-experimental designs, two were classified as quasi-experimental, and four were classified as experimental studies (references reported in Table 2; study details reported in Supplementary Material 3).

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

Non-quantitative evidence: A total of 12 non-quantitative articles were included, 9 opinions[23,32,57-63], and 3 reviews[26-28]

Opinion: Two of the nine opinion papers reported a clinical trial quality event that served as a major inflection point in thinking and approach.[57,58] These articles reported NIH tightening the clinical trials monitoring requirements in response to the finding of fraud in the National Surgical Adjuvant Breast and Bowel Project (NSABP) study group.[57,58] At the time, clinical trial monitoring practices, including SDV, varied considerably across multicenter studies funded by the National Institutes of Health (NIH).[58] Six opinion articles[23,32,34,59-61] provided guidance on different aspects of SDV or practice recommendations to improve the SDV process, multiple of them with the purpose of prompting change. One provided insight into the cost of SDV[63].

Reviews: three literature reviews were included [26-28] one review assessed SDV empirically[27] (2021 Klatte) adopted the Cochrane methodology which included only prospective empirical studies. The two remaining included reviews[26,28] focused on SDV practices. Houston et al. (2018) reported a wide range of SDV methods existing in practice with no best practice evident.[26] Similarly, Ward's review documented the absence of methodological guidelines for SDV, the wide variety in reported practice, and the lack of empirical evidence regarding the impact of SDV[28] , which stood out as the earliest identified graduate research presented as a thesis in a master's degree program.

Discussion

Critical appraisal within sources of evidence: In comparison to the Cochrane review, which included only prospective empirical studies, this review includes a much broader range of evidence from retrospective empirical studies, other reviews, and opinion papers. Though we did not explicitly assign a strength of evidence or risk of bias rating, we fully acknowledge that the retrospective, descriptive, and qualitative papers provide inherently weaker evidence due to threats to internal validity and risk of bias. For example, with one exception, there were no measures of SDV quality reported prior to broadened interest in RBM. However, the two-dimensional framework for evaluating SDV quality (Table 2) would not have emerged without considering the outcomes measured in the retrospective studies. As a scoping review, these sources were included to provide a comprehensive compilation of the available literature and

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

findings across the spectrum of research designs while at the same time recognizing the difference in evidence strength.

The recent systematic literature reviews for Good Clinical Data Management Practices also found scant Cochrane-strength evidence for practice recommendations.[65-69] In these reviews, randomized, controlled experiments of operational processes were few and far between. It may be that this standard is not a feasible expectation for evidence-supporting practice in clinical research operations. On the other hand, it still seems ironic that the methods from which strong evidence is generated are not themselves held to that same standard. Without such a standard, some may interpret the literature regarding SDV accuracy as sufficient to recommend against SDV, while others looking at the same literature will refrain from such recommendations.

Results of individual sources of evidence and synthesis of results: The included quantitative studies exhibited significant heterogeneity along multiple dimensions. We identified four reported aspects of SDV quality (rows in Table 3): (1) source data availability and access, (2) data quality, (3) study process fidelity, and (4) SDV process fidelity. Furthermore, reports for the SDV quality domains naturally fell into three main categories (columns in Table 3): (1) SDV quality measures, such as rates of findings, missing data, or process problems, (2) the impact of SDV on decreasing future problems, and (3) the impact of SDV on study results. One of the included articles assessed SDV quality on all three levels.[53] Three studies reported a measure of source data availability and access[7,51,56] (Table 3). The data quality aspect (domain) of SDV quality (second row in Table 3) was reported as data discrepancies and errors. Six articles reported counts or rates of data discrepancies[7,8,31,47,48,52], while seven reported measures of errors[33,47,49,51,53,55,56]. An error is a discrepancy with the truth, i.e., an incorrect data value. A discrepancy, on the other hand, is a difference between two data values where either or both could be in error. Discrepancies are often reported when the correct value cannot be determined or was not determined. In the included articles, some reported database changes were made after investigating and resolving discrepancies; we counted these as reports of errors.

The third and most frequently reported aspect of SDV quality was the ability to detect or propensity to miss study process errors; we refer to this as study process fidelity (Table 3). The included articles varied greatly concerning the study processes for which SDV quality was reported. Eight articles reported SDV quality with respect to the detection of safety issues[37,38,47-49,51,54,55]. One reported that SDV improved safety process fidelity over time[37]. And four reported

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

the impact of SDV on safety-related study results[47,48,51,54]. Eligibility, Informed consent, and protocol adherence were tied as the second most frequently reported aspects of study process fidelity (Table 3). These aspects encompass the EMA (2017) integrated inspection report categories used to summarize and evaluate the potential implications of major or critical findings, "the impact on the integrity of the trial data, the rights, wellbeing, and safety of the subjects, the compliance of the trial with GCP (including ethical principles) …" [70]. Within each SDV quality aspect, the included articles varied in whether and, if so, how the impact of SDV was evaluated. All sixteen included quantitative articles reported counts or rates of items detected or missed by SDV (Table 3, Figure 3). Eleven[7,8,31,33,47-49,52,53,55,56] of the sixteen quantitative articles reported measures of data discrepancies or errors. Two articles[37,53] evaluated the impact of SDV on quality improvement over time within a study, whereas nine[31,33,47,48,51-54,56] evaluated the impact of SDV findings on study results (Table 3, Figure 3). However, five[31,33,52,54,56] of the articles reporting the impact of SDV findings on study results did so qualitatively, for example, stating that the errors occurred evenly across treatment groups, that the errors occurred in non-critical variables, or that the frequency or extent of the errors was too small to have impacted the analysis. The remaining four[47,48,51,53] of the articles reporting the impact of SDV findings on study results did so by comparing analyses before and after error correction.

In addition to variability in the aspects of the SDV quality measured and the level at which they were reported (Table 3), the included articles exhibited just as much variability in the context in which they were measured. For example, seven assessments were done in industry clinical trials versus nine in investigator-initiated studies (Figure 3). The included articles were heterogeneous with respect to the SDV process for which the quality was measured. The most common SDV process variants assessed included remote, targeted, and varied extents of reduced SDV (Figure 3, Supplementary Material 3). All reported SDV quality assessments were relative, for example, comparing the number of SDV findings missed by one method that was subsequently detected by another, or reporting the portion of SDV findings that resulted in database changes. The reported assessments differed in how discrepancies or errors were identified, i.e., which two SDV processes were compared to identify items missed by one of the methods but detected by the other. The assessments differed in whether these discrepancies were verified. Multiple included studies reported the number of data discrepancies detected by SDV,

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

while others reported those missed by SDV. One study[37] comprehensively reported the number of discrepancies detected and missed by SDV but did so for only one parameter, informed consent (Supplementary Material 3). Further limiting the quantitative synthesis, the unit of analysis was inconsistent across included studies and included counts of individual findings, rates of findings per patient, rates of findings per site, or proportions of patients or sites with one or more findings. Additionally, some studies reported only major findings, while others reported all findings. These differences are detailed in Supplementary Material 3. Although multiple studies quantified one or more aspects of SDV accuracy, no article reported a comprehensive measure of SDV accuracy. No study reported the absolute rate of errors, i.e., detected against the gold standard of truth. One study[52] declared 100% SDV to be the gold standard but did not compute accuracy measures (sensitivity and specificity) against it.

The search strategy broadly encompassed articles focused on SDV as well as those focused on clinical study monitoring; relevant reports of SDV quality were found in both types of articles. For example, a study comparing 100% SDV *versus* reduced SDV that quantified items missed by 100% SDV provided an assessment of SDV quality. Similarly, a study comparing triggered *versus* non-triggered on-site visits that quantified items missed by SDV provided an assessment of SDV quality.

The majority of included articles reported SDV quality in the context of comparing RBM (including targeted, remote, or reduced SDV) to traditional monitoring approaches usually characterized by more extensive SDV. The heterogeneity, such as differences in the amount, timing, or frequency of SDV, in articles included in the quantitative synthesis (Table 3 and Supplementary Material 3) means that the assessments of SDV quality are not comparable; the methodological heterogeneity limited this synthesis to a scoping review.

The evaluative studies identified and included in this review are varied in terms of which domains of SDV quality are reported for SDV (rows in Table 3). Reported SDV quality domains included accessibility of source data, data error rates, GCP or protocol deviations (often called monitoring findings), and audit-identified deviations in the SDV process. Further, these SDV quality domains were reported at multiple levels, including accessibility of source data needed to identify errors, rates of identified errors, effectiveness at preventing future errors, and impact of

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

the identified errors on study outcomes (columns in Table 3). SDV likely has utility on each level, however, reports at the prevention and study results levels were few compared to counts or rates of findings. A comprehensive evaluation of SDV quality would include the rows and columns in Table 3 as well as items both identified and missed by SDV.

SDV accuracy cannot be calculated without a gold standard and enumeration of true positives, false positives, true negatives, and false negatives. One study[52] declared 100% on-site SDV as the gold standard; however, the authors acknowledge that errors remain in 100% SDV'd data and did not report accuracy measures of sensitivity and specificity. Thus, we did not find a quantitative report of absolute SDV accuracy (sensitivity and specificity).

Related findings: There is a good number of reviews and surveys. Despite the fact that they didn't meet the inclusion criteria, they reported related beneficial results, and we opted to present them. Seven relevant reviews[18,26,27,71-73] were identified by the search (Figure 4). Three of the identified reviews met the inclusion criteria (bottom row in Figure 4). The reviews differed in their scope, with the three included reviews specifically addressing SDV and the four excluded reviews focusing more broadly on clinical trial monitoring or other aspects thereof (Figure 4). Four of the reviews focused on methods; two included only empirical assessment of monitoring or SDV outcomes, and one review, Olsen et al. (2016), included both methods and empirical results (Figure 4). While the Ward (2013) survey did not meet inclusion criteria, the systematic review portion of this work did and is included (Figure 4). Only the three reviews, including Ward (2013), specifically addressed SDV were included in this review.

The Cochrane review by Klatte et al. (2021) included eight prospective empirical studies that compared different monitoring strategies.[27] Five of them are also included in our review. Overall, the Cochrane review concluded with moderate certainty that "risk-based monitoring is not inferior to extensive on-site monitoring with respect to critical and major monitoring findings in clinical trials" and noted that "more high-quality monitoring studies that measure effects on all outcomes specified in this review are necessary to draw more reliable conclusions". The Cochrane review did not directly address the accuracy of the SDV process itself with respect to the identification of errant data.

Eight surveys [20,28,74-79] were identified through the search, all but one [28] were excluded due to minimal focus on SDV quality. However, because each survey contained one or more SDV-

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

relevant questions, we have listed them in Supplementary Material 4 for completeness. Across the surveys, perceptions varied widely with respect to the impact of SDV or monitoring on data quality. Collectively, the survey work indicates variability in SDV frequency, amount, and methods similar to that seen in the quantitative, included articles and in the included reviews. Multiple articles reporting survey results called for additional research on SDV methods, on the impact of SDV, and more specific guidelines for methods, including the amount of SDV. Though their message is weakened by the limited generalizability of each survey and the significant changes in context and practice over the almost 30-year span over which survey results were reported.

Evidence throughout the literature, from opinion to experimental studies, supports the ability of SDV to identify unreported events. While reports of SDV-identified significant or systematic findings certainly exist,[47,48,51,52] the included articles reporting the impact of missed-event-type findings on study results indicated no significant impact.[27,37,51] Similarly, multiple included studies concluded a lack of impact of SDV-identified data errors on study results[31,48,51]. However, others found SDV-identified data discrepancies or errors impactful on one or more study analyses[7,53]. These apparent differences may reflect differences in the SDV quality domains assessed, differences in measurement methods, differences in the SDV processes themselves, or differences in other aspects of data collection and processing. For example, one of the included studies assessed the type, frequency, and impact of data errors on an observational study and concluded that data errors identified through SDV would have otherwise impacted the study results[53]. However, few upstream data quality control measures were in place. Similarly, two articles concluding no impact of SDV-identified data error measured SDV error for query-clean data and data collected on structured site worksheets (rather than abstracted from medical records) – the upstream data quality control measures described would have likely significantly decreased the number of errors remaining to be detected by SDV. The many-faceted heterogeneity precludes drawing conclusions.

With rising cost pressure, a wide variety of options continue to be explored. The Cochrane review concludes, albeit on what the authors deem to be moderate to low-quality evidence, that there is likely no difference between the different monitoring (and SDV) strategies tested in the five comparisons assessed by the review. We posit that this could be due to the likely high error rate of SDV itself, i.e., less of an error-prone inspection may not yield markedly worse overall

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

quality. One of the most convincing studies comparing the outgoing error rate from RBM to traditional monitoring concluded the same, noninferiority.[80]

Remote access to Electronic Health Records (EHRs) for monitoring and the ability to extract data directly from them greatly impact study monitoring and SDV processes. The justification for pursuing such EHR-to-eCRF data collection includes increasing data quality by decreasing manual medical record abstraction and decreasing the data collection burden at sites and Sponsors. In this case, SDV can be performed by computationally confirming that study data match the EHR source or may obviated altogether as mentioned in the FDA eSource guidance.[81] Eliminating manual SDV for EHR-to-eCRF data would likely reduce burden and cost. Given advances in technology and permissible regulatory guidance, available innovation will likely be applied to SDV, such as (1) establishing traceability back to the source that can be computationally traversed to demonstrate that the final data exactly reflects the medical record source and (2) defining a certified copy in the context of data extracted from EHRs such that sponsors and regulators are guaranteed that the final data are a replica of the source. These cannot be accomplished with manual SDV.

Limitations

Though performed systematically, our search could have missed an important article. The database searches were conducted over a 5-month period which may result in differences (though minimal) of returned articles. The methodological variation observed in the included papers was huge, making it difficult to draw conclusions from the body of the literature. For example, no two reported instances of reduced SDV were exactly alike. Similarly, the data processing and quality control applied to data prior to SDV were often not completely described. Significant differences were observed in available descriptions of upstream data processing, such as performing SDV against the medical record versus against structured site source worksheets. The contributions of SDV *versus* Source Document Review (SDR, defined in Box 1) often could not be distinguished from one another. SDV includes reading the source record to ensure that there were no omissions in the study data which includes some amount of SDR. In fact, multiple included articles remarked that SDV was conducted by reading the full source. For this reason, findings attributed to SDV may have also been found or maybe equally findable through SDR. Additionally, the variability in the terminology used in the literature could have easily led to

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

inadvertent misclassification of studies with respect to the SDV process measured, the measures used, and the methods by which they were obtained.

Conclusions

The review exposes significant heterogeneity in the SDV processes measured, the measures used, and the methods by which they were obtained. Though multiple studies quantified one or more aspects of SDV quality, we did not find an article that reported an assessment of SDV quality covering the full set of domains and levels identified in the included articles. Accuracy is not among the reported measures of SDV quality. Either of the heterogeneity or the absence of accuracy measures alone is sufficient to preclude reporting, much less comparing, measurements of SDV quality found in the literature.

Due to the likely context sensitivity of SDV, QbD (Quality by Design) should be applied to ensure that new SDV approaches will not adversely impact human safety and research results. Additional research is needed to develop methods of designing study processes capable of delivering the necessary outgoing quality. Estimates of process and inspection accuracy are needed to support prospective process design.

The variability and error associated with SDV as a manual process suggest opportunities for improvement through advances such as automation and decision support.

Funding

This work was partially supported by an Innovation in Regulatory Science Award from the Burroughs Wellcome Fund and the Clinical and Translational Science Award from the National Center for Advancing Translational Science, a component of the National Institutes of Health, to the University of Texas Health Science Center at San Antonio.

Author's Contributions:

All authors participated in the collection, analysis of data, and interpretation of results. In addition to participating and commenting on the manuscript drafted by M. Hamidi, E. and M. Zozus. All authors approved the final version for publication.

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

References

1. Getz KA, Campo RA. New Benchmarks Characterizing Growth in Protocol Design Complexity. *Ther Innov Regul Sci*. 2017;
2. Sung NS, Crowley WF, Jr., Genel M, et al. Central challenges facing the national clinical research enterprise. *JAMA*. 2003/03/14 2003;289(10):1278-87.
3. Eisenstein EL, Lemons PW, 2nd, Tardiff BE, Schulman KA, Jolly MK, Califf RM. Reducing the costs of phase III cardiovascular clinical trials. *Am Heart J*. 2005/05/03 2005;149(3):482-8.
4. Eisenstein EL, Collins R, Cracknell BS, al. e. Sensible approaches for reducing clinical trial costs. *Clin Trials*. 2008/02/20 2008;5(1):75-84.
5. Getz KA, Wenger J, Campo RA, Seguine ES, Kaitin KI. Assessing the impact of protocol design changes on clinical trial performance. *Am J Ther*. 2008/09/23 2008;15(5):450-7.
6. E6(R2) Good Clinical Practice: Integrated Addendum to ICH E6(R1) (2018).
7. Mealer M, Kittelson J, Thompson BT, al. e. Remote source document verification in two national clinical trials networks: a pilot study. *PLoS One*. 2013/12/19 2013;8(12)
8. Trancelerate. Position paper: risk-based monitoring methodology. Transcelerate Biopharma Inc.; 2013. p. 30.
9. U.S. Food and Drug Administration. Guidance for industry oversight of clinical investigations — a risk-based approach to monitoring 2013.
10. Adachi K, Shirase M, Kimura Y, Kuboki Y, Yoshino T. What and how will the risk-based approach to monitoring change? Survey of RBM in medical institutions. *Journal of the Society for Clinical Data Management*. 2022;2(2)doi:https://doi.org/10.47912/jscdm.18
11. European Medicines Agency (EMA). Reflection paper on risk based quality management in clinical trials. 2013.
12. U.S. Food and Drug Administration. Guidance for industry Q8(R2) pharmaceutical development. 2009.
13. Austin CP. Opportunities and challenges in translational science. *Clin Transl Sci*. Sep 2021;14(5):1629-1647. doi:10.1111/cts.13055
14. DiMasi JA, Grabowski HG, Hansen RW. Innovation in the pharmaceutical industry: New estimates of R&D costs. *J Health Econ*. 2016/03/02 2016;47:20-33.

15. Moore TJ, Zhang H, Anderson G, Alexander GC. Estimated Costs of Pivotal Trials for Novel Therapeutic Agents Approved by the US Food and Drug Administration, 2015-2016. *JAMA Intern Med*. Nov 1 2018;178(11):1451-1457. doi:10.1001/jamainternmed.2018.3931

16. Scannell JW, Blanckley A, Boldon H, Warrington B. Diagnosing the decline in pharmaceutical R&D efficiency. *Nat Rev Drug Discov*. Mar 1 2012;11(3):191-200. doi:10.1038/nrd3681

17. Breslauer C. Could 100% Source Document Verification Become a Risk in a Fixed-Unit Price Environment? *Monitor*. 2006:43-47.

18. Olsen R, Bihlet AR, Kalakou F, Andersen JR. The impact of clinical trial monitoring approaches on data integrity and cost--a review of current literature. *Eur J Clin Pharmacol*. Apr 2016;72(4):399-412. doi:10.1007/s00228-015-2004-y

19. Institute of Medicine. *Assuring Data Quality and Validity in Clinical Trials for Regulatory Decision Making: Workshop Report*. Roundtable on Research and Development of Drugs, Biologics, and Medical Devices, Division of Health Sciences Policy. National Academy Press; 1999:88.

20. Sandra Funning, Anders Grahnen, Karin Eriksson, Asa Kettis-Linblad. Quality assurance within the scope of Good Clinical Practice (GCP) - what is the cost of GCP-related activities? A survey within the Swedish association of the pharmaceutical industry (LIF)'s members. *Quality Assurance Journal*. 2009;12:3-7. doi:10.1002/qaj.433

21. Denis Gibbs. For Debate: 250th anniversary of source document verification. *BMJ*. 28 September 1996;313(798)doi:https://doi.org/10.1136/bmj.313.7060.798

22. Zozus MN, Pieper C, Johnson CM, et al. Factors Affecting Accuracy of Data Abstracted from Medical Records. *PLoS One*. 2015/10/21 2015;10(10):e0138649.

23. Manasco P, Bhatt DL. Evaluating the evaluators - Developing evidence of quality oversight effectiveness for clinical trial monitoring: Source data verification, source data review, statistical monitoring, key risk indicators, and direct measure of high risk errors. *Contemp Clin Trials*. Jun 2022;117:106764. doi:10.1016/j.cct.2022.106764

24. Nielsen E, Hyder D, Deng C. A Data-Driven Approach to Risk-Based Source Data Verification. *Ther Innov Regul Sci*. Mar 2014;48(2):173-180. doi:10.1177/2168479013496245

25. CG Drury, JG Fox. *Human reliability in quality control*. Taylor & Francis Ltd.; 1975:315.

26. Houston L, Probst Y, Martin A. Assessing data quality and the variability of source data verification auditing methods in clinical research settings. *J Biomed Inform*. Jul 2018;83:25-32. doi:10.1016/j.jbi.2018.05.010

27. Klatte K, Pauli-Magnus C, Love SB, et al. Monitoring strategies for clinical intervention studies. *Cochrane Database Syst Rev*. Dec 8 2021;doi:10.1002/14651858.MR000051.pub2

28. Ward R. *Examining Methods and practices of source data Verification in Canadian critical care randomized controlled trials*. Thesis submitted to the Faculty of Graduate and Postdoctoral Studies in Partial fulfillment of the requirements for the Master of Science degree in Epidemiology. 2013. https://ruor.uottawa.ca/bitstream/10393/23974/1/Ward_Roxanne_2013_thesis.pdf

29. Megaw ED. Review of : "Human Reliability in Quality Control" Edited by C. G. DRURY and J. G. Fox. (London: Taylor & Francis Ltd., 1975.) [Pp. xii+316.] £7 00. *Ergonomics*. 1976/09/01 1976;19(5):649-650. doi:10.1080/00140137608931579

30. Tantsyura V, Dunn IM, Waters J, et al. Extended Risk-Based Monitoring Model, On-Demand Query-Driven Source Data Verification, and Their Economic Impact on Clinical Trial Operations. *Ther Innov Regul Sci*. Jan 2016;50(1):115-122. doi:10.1177/2168479015596020

31. Andersen JR, Byrjalsen I, Bihlet A, et al. Impact of source data verification on data quality in clinical trials: an empirical post hoc analysis of three phase 3 randomized clinical trials. *Br J Clin Pharmacol*. Apr 2014;79(4):660-8. doi:10.1111/bcp.12531

32. Bernard J de Vries. SDV in good clinical trial practice. *Good Clinical Practice Journal*. 1996;3(1):15-17.

33. Jules T Mitchel, Timothy Cho, Dean A Gittleman, et al. Time to change the clinical trial monitoring paradigm. *Applied Clinical Trials*. January 17 2014;

34. Tantsyura V, Dunn IM, Fendt K, Kim YJ, Waters J, Mitchel J. Risk-Based Monitoring: A Closer Statistical Look at Source Document Verification, Queries, Study Size Effects, and Data Quality. *Ther Innov Regul Sci*. Nov 2015;49(6):903-910. doi:10.1177/2168479015586001

35. Houston L, Yu P, Martin A, Probst Y. Clinical researchers' lived experiences with data quality monitoring in clinical trials: a qualitative study. *BMC Med Res Methodol*. Sep 20 2021;21(1):187. doi:10.1186/s12874-021-01385-9

36. Brosteanu O, Houben P, Ihrig K, et al. Risk analysis and risk adapted on-site monitoring in noncommercial clinical trials. *Clin Trials*. Dec 2009;6(6):585-96. doi:10.1177/1740774509347398

37. Brosteanu O, Schwarz G, Houben P, et al. Risk-adapted monitoring is not inferior to extensive on-site monitoring: Results of the ADAMON cluster-randomised study. *Clin Trials*. Dec 2017;14(6):584-596. doi:10.1177/1740774517724165

38. Kim S, Kim Y, Hong Y, et al. Feasibility of a Hybrid Risk-Adapted Monitoring System in Investigator-Sponsored Trials in Cancer. *Ther Innov Regul Sci*. Jan 2021;55(1):180-189. doi:10.1007/s43441-020-00204-5

39. Chakravarthy R, Cotter K, DiMasi J, Milne CP, Wendel N. Public- and Private-Sector Contributions to the Research and Development of the Most Transformational Drugs in the Past 25 Years: From Theory to Therapy. *Ther Innov Regul Sci*. Nov 2016;50(6):759-768. doi:10.1177/2168479016648730

40. International Council for Harmonization (ICH). *E6(R2) good clinical practice: integrated addendum to ICH E6(R2) guidance for industry*. 2018. March. https://www.fda.gov/downloads/Drugs/Guidances/UCM464506.pdf

41. Clinical Trials Transformation Initiative (CITTI). Quality By Design (QBD) Toolkit. Accessed September 15, 2022. https://ctti-clinicaltrials.org/our-work/quality/qbd-quality-by-design-toolkit/

42. eClinical Forum. Risk-based approaches - best practices for ensuring clinical trial data quality. 2013. Accessed September 15, 2022. https://eclinicalforum.org/downloads/risk-based-approaches-best-practices-for-ensuring-clinical-data-quality

43. Medicines & Healthcare products Regulatory Agency (MHRA). Risk-adapted approach to clinical trials and risk assessments. In: UK DoHSC-, editor.: GOV.UK; 2022.

44. Stansbury N, Barnes B, Adams A, et al. Risk-Based Monitoring in Clinical Trials: Increased Adoption Throughout 2020. *Ther Innov Regul Sci*. May 2022;56(3):415-422. doi:10.1007/s43441-022-00387-z

45. Procedure for reporting of GCP inspections requested by the Committee for Medicinal Products for Human Use (2017).

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

46. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of internal medicine*. 2018;169(7):467-473. doi:10.7326/M18-0850

47. Maruszewski B, Lacour-Gayet F, Monro JL, Keogh BE, Tobota Z, Kansy A. An attempt at data verification in the EACTS Congenital Database. *Eur J Cardiothorac Surg*. Sep 2005;28(3):400-4; discussion 405-6. doi:10.1016/j.ejcts.2005.03.051

48. Tudur Smith C, Stocken DD, Dunn J, et al. The value of source data verification in a cancer clinical trial. *PLoS One*. 2012;7(12):e51623. doi:10.1371/journal.pone.0051623

49. Sheetz N, Wilson B, Benedict J, et al. Evaluating Source Data Verification as a Quality Control Measure in Clinical Trials. *Ther Innov Regul Sci*. Nov 2014;48(6):671-680. doi:10.1177/2168479014554400

50. Agrafiotis DK, Lobanov VS, Farnum MA, et al. Risk-based Monitoring of Clinical Trials: An Integrative Approach. *Clin Ther*. Jul 2018;40(7):1204-1212. doi:10.1016/j.clinthera.2018.04.020

51. Embleton-Thirsk A, Deane E, Townsend S, et al. Impact of retrospective data verification to prepare the ICON6 trial for use in a marketing authorization application. *Clin Trials*. Oct 2019;16(5):502-511. doi:10.1177/1740774519862528

52. Fougerou-Leurent C, Laviolle B, Tual C, et al. Impact of a targeted monitoring on data-quality and data-management workload of randomized controlled trials: A prospective comparative study. *Br J Clin Pharmacol*. Dec 2019;85(12):2784-2792. doi:10.1111/bcp.14108

53. Giganti MJ, Shepherd BE, Caro-Vega Y, et al. The impact of data quality and source data verification on epidemiologic inference: a practical application using HIV observational data. *BMC Public Health*. Dec 30 2019;19(1):1748. doi:10.1186/s12889-019-8105-2

54. Wyman Engen N, Huppler Hullsiek K, Belloso WH, et al. A randomized evaluation of on-site monitoring nested in a multinational randomized trial. *Clin Trials*. Feb 2020;17(1):3-14. doi:10.1177/1740774519881616

55. Kondo H, Kamiyoshihara T, Fujisawa K, et al. Evaluation of Data Errors and Monitoring Activities in a Trial in Japan Using a Risk-Based Approach Including Central Monitoring and Site Risk Assessment. *Ther Innov Regul Sci*. Jul 2021;55(4):841-849. doi:10.1007/s43441-021-00286-9

56. Yamada O, Chiu SW, Takata M, et al. Clinical trial monitoring effectiveness: Remote risk-based monitoring versus on-site monitoring with 100% source data verification. *Clin Trials*. Apr 2021;18(2):158-167. doi:10.1177/1740774520971254

57. Seachrist L. Scientific misconduct. NIH tightens clinical trials monitoring. *Science*. Apr 22 1994;264(5158):499. doi:10.1126/science.8160006

58. Cohen J. Clinical trial monitoring: hit or miss? *Science*. Jun 10 1994;264(5165):1534-7. doi:10.1126/science.8202707

59. Malcolm L Schuyl, Thim Engel. A review of the source document verification process in clinical trials. *Drug Information Association Journal*. 1999;33(3)

60. M. Lorstad. Data quality of the clinical trial process - costly regulatory compliance at the expense of scientific proficiency *The Quality Assurance Journal*. 2004;8(177-182)doi:10.1002/qaj.288

61. Duda SN, Wehbe FH, Gadd CS. Desiderata for a computer-assisted audit tool for clinical data source verification audits. *Stud Health Technol Inform*. 2010;160(Pt 2):894-8.

62. Tantsyura V. Risk-based source data verification approaches: prons and cons. *Drug Information Association Journal*. 7/19/2010 2010;44(0092-8615/2010):745-756.

63. K. Korieth. The high cost and questionable impact of 100% SDV. *The Central Watch Monthly*. February 2011 2011;18(02)

64. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *PLoS medicine*. 2021;18(3):e1003583-e1003583. doi:10.1371/JOURNAL.PMED.1003583

65. Hills K, Bartlett T, Leconte I, Zozus MN. CRF Completion Guidelines (CCGs). *Data Basics Society for Clinical Data Management*. Spring 2020;26(1)(Spring):33-55.

66. Lebedys E, Famatiga-Fay C, Bhatkar P, Johnson D, Viswanathan G, Zozus MN. Good clinical data management practices Data Management Plan (DMP) chapter. Data Basics, 26:1 Spring 2020: Society for Clinical Data Management; 2020. p. 74-100.

67. Eade D, Pestronk M, Russo R, et al. Web-based Electronic Data Capture (EDC) study implementation and start-up. *Journal of the Society for Clinical Data Management*. March 12 2021;1(1)

68. Montano O, Johnson D, Muthanna M, et al. Electronic Data Capture (EDC) – study conduct, maintenance and closeout. *Journal of the Society for Clinical Data Management*. March 12 2021;1(1)

69. Pestronk M, Johnson D, Muthanna M, et al. Electronic Data Capture (EDC) selecting an EDC system. *Journal of the Society for Clinical Data Management*. 2021;1(1)

70. European Medicines Agency (EMA). Procedure for reporting of GCP inspections requested by the Committee for Medicinal Products for Human Use (CHMP). 2017.

71. Hurley C, Shiely F, Power J, et al. Risk based monitoring (RBM) tools for clinical trials: A systematic review. *Contemp Clin Trials*. Nov 2016;51:15-27. doi:10.1016/j.cct.2016.09.003

72. Bakobaki J, Joffe N, Burdett S, Tierney J, Meredith S, Stenning S. A systematic search for reports of site monitoring technique comparisons in clinical trials. *Clin Trials*. Dec 2012;9(6):777-80. doi:10.1177/1740774512458993

73. Macefield RC, Beswick AD, Blazeby JM, Lane JA. A systematic review of on-site monitoring methods for health-care randomised controlled trials. *Clin Trials*. Feb 2013;10(1):104-24. doi:10.1177/1740774512467405

74. Odette Jochems JJ, Neil J Mountain, David R Hutchinson Source data verification in the Netherlands, Belgium and the UK: results of a survey to establish the current pharmaceutical industry approach. Survey. *European Journal of Clinical Research* 8 June 1993 1993 (4):45-48.

75. Hurley C, Sinnott C, Clarke M, et al. Perceived barriers and facilitators to Risk Based Monitoring in academic-led clinical trials: a mixed methods study. *Trials*. Sep 11 2017;18(1):423. doi:10.1186/s13063-017-2148-4

76. Houston L, Probst Y, Yu P, Martin A. Exploring Data Quality Management within Clinical Trials. *Appl Clin Inform*. Jan 2018;9(1):72-81. doi:10.1055/s-0037-1621702

77. Morrison BW, Cochran CJ, White JG, et al. Monitoring the quality of conduct of clinical trials: a survey of current practices. *Clin Trials*. Jun 2011;8(3):342-9. doi:10.1177/1740774511402703

78. Michael R Hamrell, Kathleen Mostek, Lyn Goldsmith. Monitoring of clinical trials - are remote activities helpful in controlling quality? *Clinical Researcher*. 2016;October 2016

79. C Kunzl, B Breuer, Y Rollinger, M Sigmund, V Kunert. RBM-An update of experiences among European CRAs *Applied Clinical Trials*. 2017. Accessed 10/19/2017. https://www.appliedclinicaltrialsonline.com/view/rbm-update-experiences-among-european-cras

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

80. Andersen JR, von Sehested C, Byrjalsen I, Popik S, B. FA, Bihlet AR. Impact of monitoring approaches on data quality in clinical trials. *Br J Clin Pharmacol*. 2022:1-11. doi:10.1111/bcp.15615

81. U.S. Food and Drug Administration. Guidance for Industry Electronic Source Data in Clinical Investigations 2013.

1.a 1.b

Figure 1: Visual Inspection Exercise

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

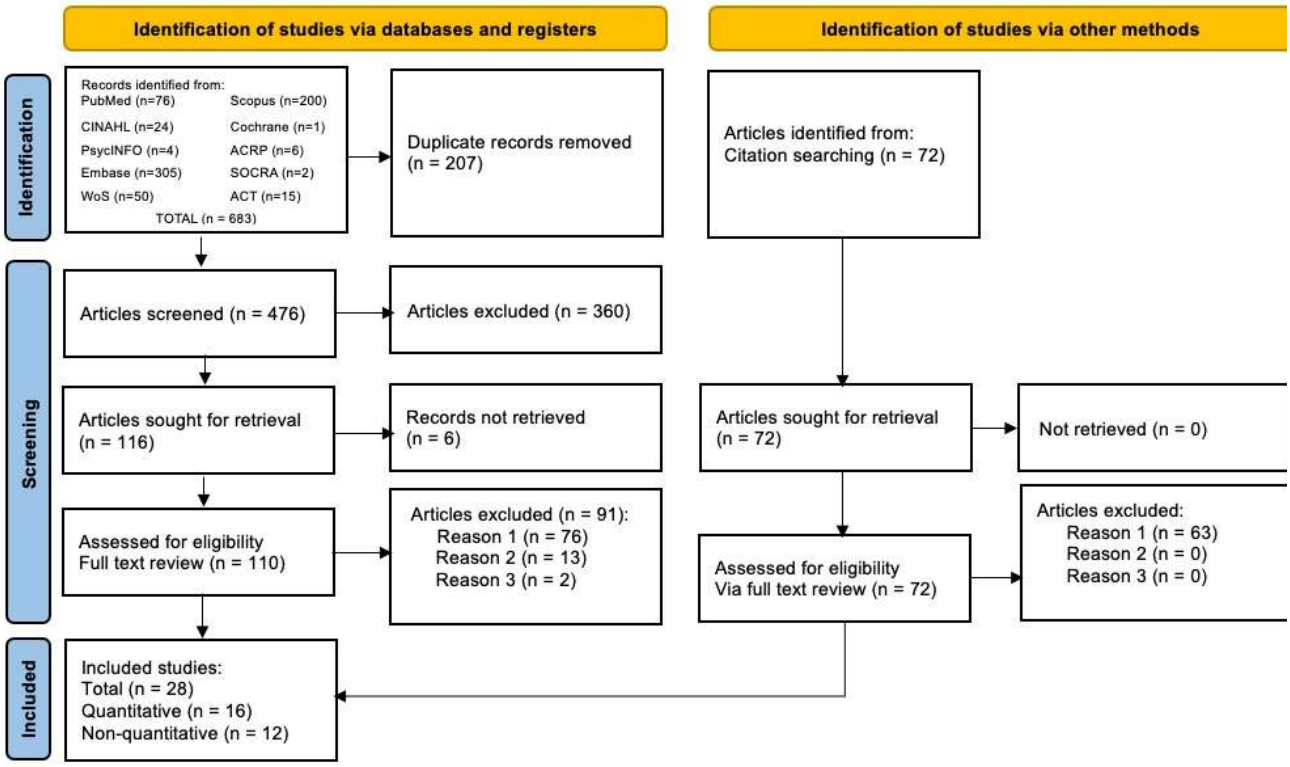

Figure 2: Flow diagram[64]

Web of Science database (WoS), American Psychological Association PsycInfo database (PsycINFO), EMBASE database (Excerpta Medica dataBASE), The Association of Clinical Research Professionals (ACRP), The Society of Clinical Research Associates (SOCRA), Applied Clinical Trials (ACT).

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

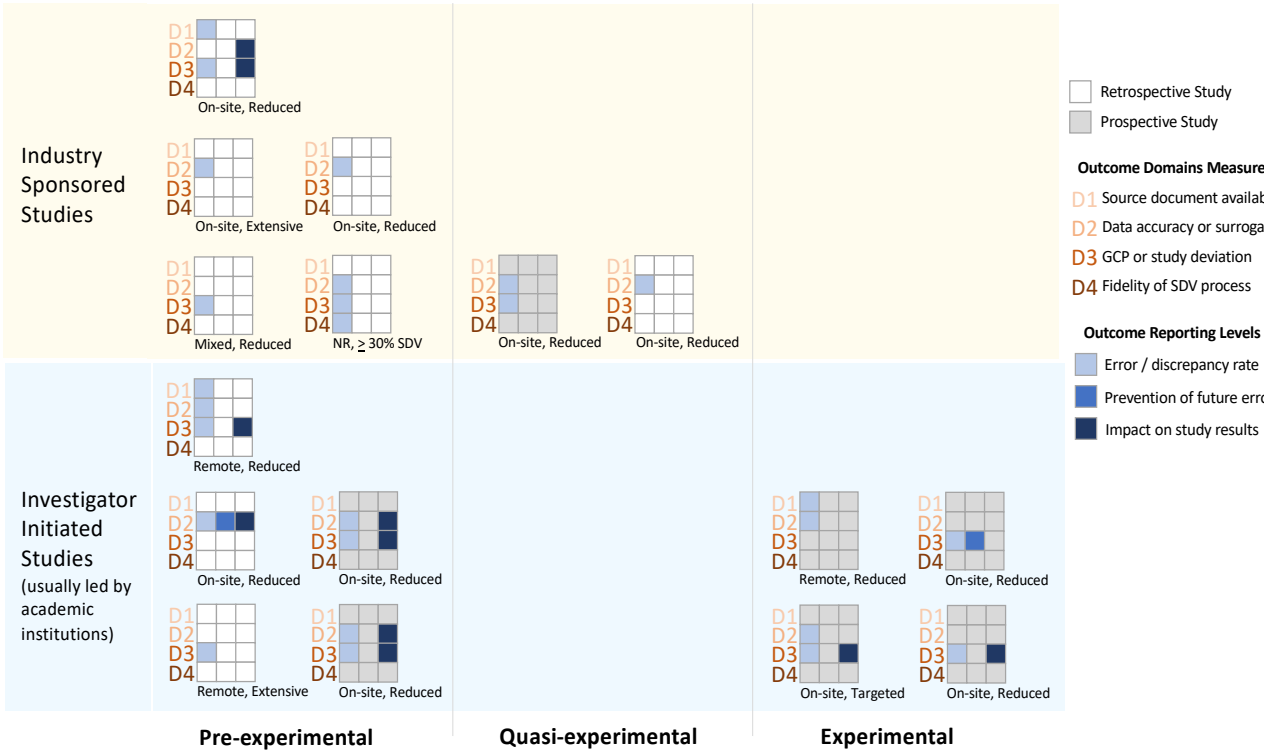

Figure 3: Heterogeneity in SDV[1] Quality Assessment

Each 3x4 grid in the figure represents SDV quality assessments reported in one included, quantitative article. The SDV methods compared are listed at the bottom of each grid, with NR signifying not reported, extensive signifying high amounts of data values undergone up to 100% SDV, and mixed signifying a combination of two or more SDV methods.

[1]Source Data Verification (SDV): the comparison of study data to their original recording to ensure that, "the reported trial data are accurate, complete, and verifiable from source documents."[6]

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

	Methods or Practices	Empirical Assessment	
Monitoring	Macefield 2013 Hurley 2016 Olsen 2016	Bakobaki 2012	**Excluded**
SDV[1]	Houston 2018 Ward 2013	Klatte 2021	**Included**

Figure 4: Categorization of Review Articles

[1]Source Data Verification (SDV): the comparison of study data to their original recording to ensure that, "the reported trial data are accurate, complete, and verifiable from source documents."[6]

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

Table 1: Information abstracted from articles

Article Metadata / Bibliographic Information	year published Author's Last name Article Title Volume issue Journal name
Inclusion / Exclusion	Included or excluded after full text review If excluded after full text review, list exclusion reasons. Group Adjudication/Final Determination
Review Assignments	Reviewer 1 name Reviewer 2 name
Information about the article or paper	Article reports Primary research results (1=affirmative, 0=negative) Article reports Secondary research results, i.e., cites/mentions results reported elsewhere (1=affirmative, 0=negative) Article reports ONLY opinion or consensus statements as the main focus (1=affirmative, 0=negative) Purpose of the paper (free text) Years over which study was conducted
Information about Research Design	Research design (Experimental, quasi-experimental, pre-exp., descriptive-quantitative, descriptive-qualitative, non-research) If experimental; Control and comparator? If experimental, Allocation method? Outcome / dependent variables (i.e., the thing that you expect to change after the exposure or intervention) and operational definitions. When with respect to (wrt.) Intervention or exposure were observations made (Before, After or Both) Mostly for experimental, quasi-experimental or pre-experimental studies, were the study Results / interpretation uncertain or clear-cut? Study Strengths. Study weaknesses.
Information about how SDV Accuracy was measured in the study	SDV* Accuracy Measured (1=affirmative) How was SDV Accuracy Measured Unit of Analysis for the SDV error rate (data field, record, form/page, visit, research subject, Abstractor or Monitor, research site, study) Number of data values assessed (error rate denominator) Number of data errors identified (error rate numerator) Accuracy Statistics Reported (agreement, chance-adjusted agreement, error rate, sensitivity, or specificity) Other aspects of SDV quality measured (free text) SDV Accuracy Results interpretation (equivocal vs. clear)
Information about the research context, e.g., type/s of studies in which the research was	Parent study NCT[1] number (if any) Parent study type (RCT[2], Registry, Correlational, PCT[3], CER[4], HSR[5], Epidemiology, other) Parent study therapeutic area Parent study context – FDA[6] vs. NIH[7] regulated

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

conducted	Research context - Sponsor-led vs. Investigator-led Type of source documents, e.g., paper charts vs. electronic source Medical Record Abstraction Quality Assurance used (Standardized Abstraction Methods, Abstraction Training, Abstraction Environment, Abstraction Process Control via re-abstraction) Description of pre-SDV[8] data processing, such as query rules ran prior to SDV (free text) Percent SDV SDV Method (Read full chart to confirm representation on study forms, confirmed data on study forms by locating value in the chart, other) Who carried out SDV (Study Coordinator, Monitor, Auditor, other)
Comments	Free text field for the reviewers to make any comments not represented in other columns
Article Acquisition	1=Obtained; 0=Un-obtainable

[1]NCT Number: The National Clinical Trial number is an identification that ClinicalTrials.gov assigns a study when it is registered.
[2]RCT: Randomized Controlled Trial. [3]PCT: Pragmatic Clinical Trial. [4]CER: Clinical Evaluation Report. [5]HSR: Health Services Research. [6]FDA: The Food and Drug Administration. [7]NIH: National Institutes of Health. [8]Source Data Verification (SDV): the comparison of study data to their original recording to ensure that, "the reported trial data are accurate, complete, and verifiable from source documents."[6]

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

Table 2: Research Design Summary for the Included Quantitative Studies

Design	**Prospective**	**Retrospective**	**Total**
Experimental	4[7,37,52,54]	0	4
Quasi-experimental	1[55]	1[31]	2
Pre-experimental	2[47,48]	8[8,33,38,49-51,53,56]	10
Total	7	9	16

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

Table 3: Aspects of SDV[1] Quality and the Level at Which They Were Reported

SDV Quality Aspects (Domains of SDV Quality)	Reports of Measures in the SDV Quality Domain	Reports of SDV Preventing Future Occurrences	Reports of SDV Impact on Study Results
(1) Source data availability and access	3[7,51,56]	-	-
(2) Data quality (DQ)*			
Discrepancies	6[7,8,31,47,48,52]	-	3[31,48,52]
Errors	6[33,47,49,53,55,56]	1[53]	2[33,51,53]
(3) Study process fidelity[2, 3]			
Human subject protection	-	-	-
IRB oversight	1[38]	-	-
Informed consent	6[37,38,52,54-56]	1[37]	1[56]
Privacy	-		1[54]
Safety	8[37,38,47-49,51,54,55]	1[37]	4[47,48,51,54]
Site study team training	-	-	-
Research subject disposition	-	-	-
Identification and screening	-	-	-
Eligibility	6[37,38,48,52,54,56]	1[37]	3[48,54,56]
Enrollment	-	-	-
Allocation/Exposure	1[38]	-	-
Retention	1[54]	-	-
Completion/Withdraw	1[54]	-	1[48]
Protocol Adherence**	4[37,48,55,56]	1[37]	5 [47,48,52,54,56]
Investigational product	1[38]	-	-
Essential documents	1[38]	-	-
Unanticipated problem identification and handling		-	-

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press

Unspecified audit findings	1[50]	-	-
Unsubstantiated data alteration or fraud	-	-	-
(4) SDV process fidelity	1[49]	-	-
Number of Distinct Articles Reporting at Each Level	**16**	**2**	**9**

[1]Source Data Verification (SDV): the comparison of study data to their original recording to ensure that, "the reported trial data are accurate, complete, and verifiable from source documents."[6]

[2]These assessments were done relative to another method rather than as a quantification of absolute errors. i.e., the studies have not measured SDV actual accuracy, instead, surrogates for accuracy were measured, such as a number of data discrepancies or errors missed relative to some other method.

[3]Includes delivery of the intervention, endpoint Assessment, regulations, guidance, and other requirements to which the protocol must comply.

https://doi.org/10.1017/cts.2024.551 Published online by Cambridge University Press