Assessing data quality and the variability of source data verification auditing methods in clinical research settings

Lauren Houston  
*University of Wollongong, lah993@uowmail.edu.au*

Yasmine Probst  
*University of Wollongong, yasmine@uow.edu.au*

Allison Martin  
*University of Wollongong, allisonh@uow.edu.au*

Follow this and additional works at: [https://ro.uow.edu.au/ihmri](https://ro.uow.edu.au/ihmri)

Part of the Medicine and Health Sciences Commons

**Recommended Citation**

Houston, Lauren; Probst, Yasmine; and Martin, Allison, "Assessing data quality and the variability of source data verification auditing methods in clinical research settings" (2018). *Illawarra Health and Medical Research Institute*. 1412.  

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au
Assessing data quality and the variability of source data verification auditing methods in clinical research settings

Abstract

Introduction: Data audits within clinical settings are extensively used as a major strategy to identify errors, monitor study operations and ensure high-quality data. However, clinical trial guidelines are non-specific in regards to recommended frequency, timing and nature of data audits. The absence of a well-defined data quality definition and method to measure error undermines the reliability of data quality assessment. This review aimed to assess the variability of source data verification (SDV) auditing methods to monitor data quality in a clinical research setting.

Material and methods: The scientific databases MEDLINE, Scopus and Science Direct were searched for English language publications, with no date limits applied. Studies were considered if they included data from a clinical trial or clinical research setting and measured and/or reported data quality using a SDV auditing method.

Results: In total 15 publications were included. The nature and extent of SDV audit methods in the articles varied widely, depending upon the complexity of the source document, type of study, variables measured (primary or secondary), data audit proportion (3-100%) and collection frequency (6-24 months). Methods for coding, classifying and calculating error were also inconsistent. Transcription errors and inexperienced personnel were the main source of reported error. Repeated SDV audits using the same dataset demonstrated ∼ 40% improvement in data accuracy and completeness over time. No description was given in regards to what determines poor data quality in clinical trials.

Conclusions: A wide range of SDV auditing methods are reported in the published literature though no uniform SDV auditing method could be determined for "best practice" in clinical trials. Published audit methodology articles are warranted for the development of a standardised SDV auditing method to monitor data quality in clinical research settings.

Disciplines
Medicine and Health Sciences

Publication Details
Assessing data quality and the variability of source data verification auditing methods in clinical research settings

Corresponding author
Lauren Houston
School of Medicine, University of Wollongong
Northfields Ave,
Wollongong, NSW 2522
E: lah993@uowmail.edu.au
M: +61 430 371 929

Author and co-authors
Ms. Lauren Houston
School of Medicine, Faculty of Science, Medicine, and Health, University of Wollongong, Wollongong NSW 2522, Australia
Illawarra Health and Medical Research Institute, University of Wollongong, NSW 2522, Australia
E: yasmine@uow.edu.au

Dr. Allison Martin
School of Medicine, Faculty of Science, Medicine, and Health, University of Wollongong, Wollongong NSW 2522, Australia
Illawarra Health and Medical Research Institute, University of Wollongong, NSW 2522, Australia
E: allisonh@uow.edu.au
Abstract

Introduction: Data audits within clinical settings are extensively used as a major strategy to identify errors, monitor study operations and ensure high-quality data. However, clinical trial guidelines are non-specific in regards to recommended frequency, timing and nature of data audits. The absence of a well-defined data quality definition and method to measure error undermines the reliability of data quality assessment. This review aimed to assess the variability of source data verification (SDV) auditing methods to monitor data quality in a clinical research setting. Material and methods: The scientific databases MEDLINE, Scopus and Science Direct were searched for English language publications, with no date limits applied. Studies were considered if they included data from a clinical trial or clinical research setting and measured and/or reported data quality using a SDV auditing method. Results: In total 15 publications were included. The nature and extent of SDV audit methods in the articles varied widely, depending upon the complexity of the source document, type of study, variables measured (primary or secondary), data audit proportion (3-100%) and collection frequency (6-24 months). Methods for coding, classifying and calculating error were also inconsistent. Transcription errors and inexperienced personnel were the main source of reported error. Repeated SDV audits using the same dataset demonstrated ~40% improvement in data accuracy and completeness over time. No description was given in regards to what determines poor data quality in clinical trials. Conclusions: A wide range of SDV auditing methods are reported in the published literature though no uniform SDV auditing method could be determined for “best practice” in clinical trials. Published audit methodology articles are warranted for the development of a standardised SDV auditing method to monitor data quality in clinical research settings.

Keywords: Source data verification, data quality, clinical trial, quality assurance, audit
1. Introduction

Clinical trials are vital to enabling a greater understanding of how interventions work in humans [1]. Therefore, it is essential that clinical trials produce accurate, complete and relevant data [2, 3]. The integral nature of good quality data and documentation practice is well accepted within the research community as conclusions and recommendations rely heavily on the outcomes of the data. As part of quality assurance practices, modifications to study practices are made to prevent errors occurring, however, no universally accepted method for measuring error rates currently exists [4]. Within the published literature, it is generally accepted that if greater than 10% of data is missing or incorrect, analysis of the data is considered to be unreliable [5, 6]. Data audits are conducted to verify that data is appropriately documented, coded and classified and may assess compliance to a protocol. To allow for this all data must be recorded on source documentation to reconstruct the trial as it happened and allow for an independent observer to confirm the data validity [7]. Without source data, audits cannot be completed and the fundamental principle of protecting participant rights, safety and well-being cannot be guaranteed. Therefore, data audits work closely with quality assurance processes by allowing on-site monitoring activities to aid study investigators to improve data quality and overall study operations. Quality assurance audits, within clinical settings, are extensively used and are a significant strategy to ensure high-quality data [4, 5, 8, 9]. However, on-site audits are infrequently published or reported. This may be due to non-specific audit methods, lack of time and funding or the publishing of audit methodology is not seen as a ‘value added’ activity [10-12]. Published methodology papers are warranted to promote routine auditing within academic and commercial clinical research settings.

In clinical trials auditors should evaluate whether the data is collected and managed in accordance with a known quality standard such as the International Conference on Harmonisation (ICH) guideline E6 on Good Clinical Practice (GCP), 1996 [13]. According to the ICH-GCP, there is a need for on-site monitoring before, during and after a clinical trial, but they do not specify the frequency and nature of such monitoring. This refers to the complete clinical record of the participant and includes the process of applying eligibility criteria and participant consent throughout the trial
until the participant completes the study and the report is published [7]. The practice
of source data verification (SDV) fulfills the ICH-GCP requirements and is a process
of comparing data collected on original source documents to data recorded on a
case report form (CRF) or electronic record [13]. Source documents are considered
the "gold standard" from which data is obtained in clinical trials [10, 14, 15]. The
United States Food and Drug Administration (FDA) released a guideline in 1988 for
Monitoring of Clinical Investigations leading to a general agreement that completing
100% SDV was required to meet the needs for high data quality and integrity [16].
These guidelines have since been withdrawn. The updated guidelines [17] promote
the use of alternative monitoring methods such as, risk-based approaches to
monitoring. Similarly, an updated ICH guideline E6(R2) was released in 2015 which
addressed audits, though the guideline still does not provide detail regarding auditing
methods and an ‘acceptable’ level of error [18]. The guidelines suggest that “when
significant noncompliance is discovered, the sponsor… should perform a root cause
analysis and implement appropriate corrective and preventative actions. If required
by applicable law or regulation the sponsor should inform the regulatory
authority(ies) when the noncompliance is a serious breach of the trial protocol or
GCP” [18]. Without a clear definition of what ‘noncompliance’ and ‘serious’ breach
refer to, it is left up to interpretation of the investigator. This leaves individual
organisations to set their own approach for data auditing, thus justifying the need for
clear and suitable systematic methods to monitor data quality.

The cost-effectiveness of SDV has been questioned; it is an expensive and labour
intensive activity [10] with estimates that monitoring site data represented 25-30% of
the total study cost [19]. Research grants seldom include funding for an audit
program [20], however, costs may be reduced by 40% by simultaneously reducing
the amount of data collected and number of on-site visits. Further, as quality
assurance improves and is reinforced throughout a study, the cost of completing
audits, in turn, decreases [10]. Nevertheless, it is recognised that effective annual
data audits on 10-20% of records are a cheaper solution [10] though the suggestion
of a 5-10% random sample of participant records has been considered adequate for
on-site auditing [21]. It is often assumed that a greater percentage of data audited is
associated with improved data quality; although there is no experimental or statistical
evidence to support this premise [22].
Thus, the aim of this review is to assess the variability of SDV auditing methods to monitor data quality in a clinical research setting. It was hypothesised that data quality variations are due to unclear SDV auditing methods existing within clinical research settings.

2. Material and methods

A literature review was conducted with reference to frameworks for establishing evidence for practice provided by the Australian Government National Health and Medical Research Council (NHMRC) [23] and followed the requirements of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist [24]. To ensure a standard and detailed method was applied the Cochrane Handbook for Systematic Reviews of Interventions was consulted for guidance [25]. Studies reviewed were assessed for quality to identify potential risk of bias using the Academy of Nutrition and Dietetics Evidence Analysis Library quality rating checklist [26]. The review addressed the research question, “Is there consistency in SDV auditing methods to monitor data quality in clinical research?”

A methodical search was conducted using MEDLINE, Scopus and Science Direct scientific databases to identify relevant articles that had assessed data quality through the use of a SDV audit. An identical search strategy was applied using the following keywords: (source data verification OR source document verification OR data monitor* OR data collect*) AND (quality control OR quality assurance) AND (medical trial OR medical record OR clinical trial OR clinical record) AND (medical audit OR clinical audit OR data audit OR audit) AND (data quality OR data integrity OR data quality improvement) AND (error* OR data). No date limits were set for the search. Publications were restricted to those published in the English language due to lack of translation resources and included only human participants. This ad-hoc literature review was undertaken to focus on clinical research trials who have implemented SDV auditing methods. Medical Subject Heading (MeSH) were deemed inappropriate due to the topic of research being an emerging area within the literature. Specific MeSH terms are used to describe articles for medical records
while this research focuses on clinical research. Potentially relevant articles were
screened by title and abstract and subsequently through full-text if required. Where
data was not immediately available in the published article, reference lists were
hand-searched and corresponding authors were contacted to clarify outcomes.

### 3. Results

A total of 802 articles were identified with 183 articles not meeting the inclusion
criteria (49=non-English and 134=non-human participants). Please refer to Figure 1
for full details.

#### 3.1 Description of studies

Ultimately, 15 relevant published articles were scrutinised and summarised in Table
1. Eight articles assessed the accuracy, completeness, and reliability of a database
[14, 15, 27-32] and seven assessed the performance and value of a SDV method [5,
6, 9, 12, 33-35]. There was substantial heterogeneity in study design, in particular
with respect to error coding, classification, and calculation. Three articles reported
data from randomised controlled trials (RCTs) [27, 28, 33]; three from comparative
cohort studies [9, 15, 29]; two from interrupted time series with no control [12, 14];
two from historical control studies [5, 6]; three from cohort studies with no
comparison [30, 31, 34]; one from a single-arm study [35] and one from a pre-
test/post-test study design [32]. All 15 articles were deemed to have sound study
design and scientific rigour overall (Additional file 1).

The types of clinical studies from the 15 included articles are summarised in Table 2.
Although no year restriction was applied, the majority of publications were published
within the past ten years (n=10) and represent studies in six different countries within
institutional (n=10) and academic research (n=5) organisations. Studies
implemented various auditing methods to suit individual policy and procedures for
study design.

#### 3.2 Methods of source data verification
The nature and extent of the SDV audit methods varied depending upon the amount of data collected and the complexity of the source document [5]. Data inconsistencies resulted from how studies recorded information and how they entered, stored and formatted the data [29]. Two articles demonstrated that health records and auditing tools differed within the same study [9, 29]. Good data quality also appeared to be related to the effectiveness of data-monitoring plans which had a functioning structure and efficiently organised data [34]. In turn, knowledge, and experience of data entry personnel involved in data management were related to the error rate [34]. Abstraction and transcription of data were identified as the steps most likely to introduce error [5, 34]. Similarly, the design of the CRF was considered important to minimise error and needs to be emphasised [34].

Three publications which implemented multiple SDV audits on the same dataset found an improvement in data quality over the duration of the study [12, 15, 29]. Completion of a re-audit reported a 50% decrease in the overall error rate. The remaining errors existed due to data existing on the source document but not entered into the database, rather than minor and major incorrect values [29]. Mphatswe et al. [12] found the level of completeness increased 38% and accuracy increased 28% when comparing the first and third audits. A re-audit of data elements was identified as suboptimal and should only be targeted for further continuous data quality interventions [15, 32].

SDV was considered to be time-consuming, expensive and not necessarily free from error [9, 28, 33, 35], however, interpretation of on-site issues provided much more than just identification of error [6]. For example, identifying whether errors were systematic allowed for potential solutions and recommendations to be introduced for overall quality improvement [5, 6]. Two publications suggested that central monitoring, such as remote SDV and risk assessment are more effective at identifying data errors when compared to traditional on-site 100% SDV, the method recommended by the ICH-GCP [9, 33]. For a precise understanding of how and where errors lie within a dataset, on-site monitoring can provide greater insight into problems and aid in identifying potential solutions for improvement, not only within the data, but the clinical study itself [6].
3.3 Variables, frequency, and amount of source data verification

Only one included publication completed traditional on-site 100% SDV [34]. All others completed a random sample of data points, CRFs, study participants or centres. Key data variables were not defined consistently but included important data items such as primary and secondary endpoints, informed consent, eligibility criteria, randomisation distribution, adverse events and safety data [5, 9, 12, 14, 27, 28, 33]. Further to this, seven publications did not state or specify if variables were critical to the outcome of the study [6, 15, 29-32, 35].

The frequency of SDV audits was stated in nine publications, in which two conducted SDV before, during and after data entry [12, 28], though one did not report on specific time intervals [12]. Time points varied (6-24 months) between publications [9, 15, 28, 31, 35] and this was due to study design and/or implementation of study interventions such as an electronic medical record (EMR) [14]. Two publications depended on completing a second audit based on the percentage of expected participant enrollments (20-30% and 70-80%) [5] as well as identifying sites that required major quality interventions [29]. However, no description was given in regards to what determined poor quality data.

The number of participants, files and centres included for each SDV audit varied and the amount of SDV varied depending on the nature and size of the data generated. Random sampling was the most common auditing method implemented, however the amount of participants (8-94%) [9, 14, 15, 27-30, 33], files (3-10%) [31, 35] or centres (11-35%) [6, 12, 32] randomly selected differed substantially. The remaining two publications implemented methods including; a specific number of files per centre, [34] and expected participant enrolment [5].

3.4 Error coding, classification, and calculation

The method used to code, classify and calculate error varied widely. Of the 14 publications that classified or coded data only eight provided clear definitions for the codes [5, 9, 12, 27, 29, 31, 34, 35]. One did not specify coding or an explanation of how error rates were calculated at all [28]. Published auditing methods were implemented in four publications, the National Cancer Institute (NCI) method for auditing cases [5, 27] and the European Organisation for the Research and
Treatment of Cancer (EORTC) [29, 34]. The total error was described in the literature by dividing total erroneous and missing points by total audited points [5, 14, 29], a standard error calculation [30, 31], and calculating an agreement rate [6, 12]. Thus, dissimilarities in the available auditing methods were a key finding demonstrating the heterogeneity in SDV auditing methods. Identifying errors through retrospective checking and data cleaning was a more successful method when compared to detection through data entry itself [28]. Reporting of error rates depended on how data was abstracted within the individual studies and results varied greatly (<1% to 71%). No systematic pattern was determined for CRF-to-database audits. Reported discrepancies had a minor impact on the primary outcome of the study, with the error reported at <1% [5, 33]. On the other hand, both source document-to-database and source document-to-CRF average error rates were much greater (~10-20%) for the majority of publications (n=9). The main source of error (0.4-14.5%) was from transcribing data from paper to electronic records [34]. Knowledge and experience of the data management and entry personnel were directly related to efficient data collection and organisation, which was linked to the percentage of error found [34].

Visual inspection, by manually checking data through SDV, allows identification of errors that fall outside predetermined values [28, 32]. The majority of publications implemented continuous extensive range, logic, and consistency checks via an electronic database [5, 6, 28, 30, 32, 33]. Errors within these articles were considerably lower when compared to those that did not implement additional checks. Errors were identified instantly and correctly resubmitted prior to auditing [33].

4. Discussion

This literature review has identified that there are limited methodological publications available on quality assurance procedures within clinical research settings. With the available published data, this review highlights the heterogeneity of SDV auditing practices and the significant variations in procedures, policies, requirements, and technologies of the audit designs used in clinical studies. As only four publications used an established SDV audit method considerable variation amongst terminology
and methodology was found [5, 27, 29, 34]. Parallel with results from this literature
review a survey conducted by the Society of Clinical Data Management reported up
to seven different methods used to calculate error rates leading to variances in the
results [4]. Therefore, it is evident there is a need to create clear definitions and
guidelines to avoid ambiguity.

This literature review identified that completing 100% SDV requires careful
consideration as it does not guarantee error-free results. Interestingly, only one
publication in this review completed traditional 100% SDV of all data points [34].
Additionally, key data for auditing was defined inconsistently throughout studies and
variations may have been subject to the study investigators personal judgement.
However, completing 100% SDV on secondary data that is not subsequently used
for analyses may not be cost-effective. This aligns with Eisenstein et al. [19] who
compared the amount of data collected and the effect this had on the cost of the
clinical trial. They found that the most efficient way to reduce trial costs, without
compromising scientific objectives, was to reduce the number of on-site monitoring
visits. Despite this, the Institute of Medicine is concerned that any discrepant data
points identified during monitoring will bring into question the reliability and validity of
the whole dataset [36]. From the outcomes of this review the authors suggest that
data audit characteristics such as conducting 100% SDV on critical endpoints and
random selection of 10% SDV of non-critical endpoints, could be a more accurate
and cost-effective method to ensure data quality. A quality improvement plan should
be implemented for a follow-up audit if the error is found to be greater than 10%, to
revise and improve site procedures. This suggested data audit method would need
to be tested against a traditional method (e.g. 100% SDV) to provide evidence for
implementation.

Although there is currently a lack of knowledge regarding alternative auditing
methods, the literature supports changing the focus from on-site SDV audits to
employing other methods to monitor clinical trials [33]. Central monitoring techniques
such as risk-based monitoring, which include identifying critical errors by a risk-
algorithm or risk-assessment model, have been recently welcomed by the research
community. Critical errors are those that are crucial to the quality of the data and
may affect the primary outcome of the study. Thus, quality assurance is determined
based on the level of risk identified and assessed on various critical data points or activities. Two large academic groups in France [37] and Germany [38] are conducting large RCTs comparing risk-based monitoring and on-site SDV and compliance with GCP. The German group recently (August 2017) published findings that risk-based monitoring utilises less than 50% of resources compared to intensive on-site monitoring, which only had a small benefit (8.2%) relative to the overall findings [39]. This study is a major step towards improving auditing methods.

Previously in the literature, other researchers have captured and overlapped essential elements of established sociotechnical models to evaluate the acceptance of new technology into a unified theoretical model [40]. The researchers’ suggest, based on the analysis of the current literature, that future research should draw on a similar approach and include the strongest elements of previously tested methods and investigate the feasibility of creating a single consolidated method for all clinical research. However, it is essential that the method allow for adaptations according to study context and data availability.

With emerging literature on technology improvements, there are new opportunities and concerns about the increase in reuse of clinical research data. The American Medical Informatics Association stressed the urgency and complexity of issues that surround the secondary use of clinical data and compiled a set of recommendations that warrant the need for standards and clear rules to define evidence [41]. Additionally, Mackenzie et al. [42] conducted a survey that found one of the primary obstacles related to integrated data repositories included data quality and standard issues. It is evident that with our improvements in clinical systems and data sharing, this issue is of great importance now more than ever. Clearer auditing methods are required to ensure high data quality is guaranteed during primary data collection and at a larger-scale for the use of clinical data in biomedical research.

The scope of this literature review was limited to the small number of articles identified by the selection of keywords. It is presumed that many clinical studies undergo audits, but do not publish their methods, as they may be considered confidential by many organisations [4]. According to the ICH-GCP, the FDA does not assess the results of audits directly and rather guides the sponsor(s) on how to conduct audits [18]. However, procedures on how to audit, what to audit, the
frequency of audits, and the form and content of audits are left to the sponsor to
design and determine. It is recommended that each country’s regulatory authority
(European Union – European Medicines Agency (EUA), Japan – Pharmaceuticals
and Medical Devices Agency (PMDA), Canada – Therapeutic Products Directorate
(TPD) and Australia – Therapeutic Goods Administration (TGA)) guide this process
but only seek audit reports when serious legal proceedings and/or GCP non-
compliance exists [18]. It is understood that the pharmaceutical industry enforces
tightly regulated monitoring activities and may have determined what methods are
optimal, though this again cannot be determined from the current published
literature.

SDV audit methods are varied because each article is assessing different primary
outcomes. The majority of the literature included for the review was considered to
have low levels of evidence according to study types [23]. Additionally, the focus on
clinical research trials may of ruled out relevant articles from the informatics and
library science literature. Limited articles took into consideration whether a SDV audit
was effective in achieving quality assurance. These limitations made it difficult to
determine and justify clinical significance in relation to acceptable data quality. The
lack of evidence in the literature regarding methods of SDV is a strong indication of
the need to conduct further research in this area. The lack of adequate detail in the
original publications may also have affected the strength of our conclusions. All
articles differed in their study size, country of origin, ethnicity, education level,
resources and method for implementing a SDV were varied. Therefore, the main
limitation of this review is being unable to compare the methods of SDV auditing
against each other. As this area of research is growing, it was the researchers’
decision to not use only MeSH terms, descriptors and concepts. Consideration is
required as the review spans a number of disciplines and limiting to MeSH terms
may in turn limits the search strategy and number of relevant articles returned.
Although MeSH terms were not incorporated in the search strategy, a stringent
systematic procedure was followed and guided by the NHMRC framework and
PRISMA checklist. Further, a quality-rating process was implemented to identify
potential bias and a range of scientific databases were utilised as recommended by
the Cochrane handbook highlighting the strength of this review.
Overall analyses of included articles suggest that the importance of data structure and the added potential for correcting errors are valuable. Future research is required focusing on the value of SDV to support trial investigators who wish to ensure data integrity, but must work within the limits of funding and resource availability. Unique and different approaches for data audit characteristics may be more appropriate and need to be taken into consideration for example; comparing RCT to observational studies. Determining an acceptable error rate is also warranted to assess the quality and monitor clinical research data. The question of “how many errors are too many?” is difficult to answer as the included variables and the robustness of the study design play a significant role in determining the amount of error.

5. Conclusion

This literature review provides an overview of the use of SDV auditing methods and the issues related to its use. The review demonstrates few articles reporting audit methods in a clinical setting. There is inconsistency in the methodology of SDV auditing in clinical research studies and a scarcity of evidence to support best practice. Therefore a gap in the literature has been identified and a need to assess SDV auditing methods in clinical studies. Based on the small number of identified publications and a lack of experimental evidence, no uniform SDV auditing method or approach can be determined for the conduct of clinical trials. The utilisation and improvement of SDV auditing methods are required. Based on the current literature recommendations for using a combination of a random sample of participants’ records (≥10%); both critical and non-critical variables and multiple audits (before, during and after) with quality improvement feedback should be included in the SDV audits. This combination is considered to be a cost-effective solution to ensure data quality, at least until further studies are conducted.
List of abbreviations

CRF: Case report form
EMR: Electronic medical record
ICH-GCP: International Conference on Harmonisation – Good Clinical Practice
NCI: National Cancer Institute
NHMRC: National Health and Medical Research Council
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SVD: Source data verification

Competing interests

The author(s) declare that they have no competing interests.

Declarations/Author contributions

This research received no specific grant from any funding agency, commercial or not-for-profit sectors. L. H. conceptualised and formulated the research question, designed the study, carried out the study, evaluated the data, drafted the initial manuscript, revised the manuscript and approved the final manuscript as submitted; Y. P. and A. M. made substantial contributions to the study design, analysis, and interpretation of the data critically reviewed the manuscript, edited and approved the final manuscript as submitted.

LEGEND/CAPTIONS

Figure 1: Search strategy PRISMA flow diagram
REFERENCES


20. Ward R: Examining methods and practices of source data verification in Canadian critical care randomised controlled trials. *University of Ottawa: Faculty of Medicine, Ottawa, Canada*, 2013.


Records identified through database searching
Search 1: MEDLINE (n=223), Search 2: Scopus (n=132), Search 3: Science Direct (n=447) (n=802)

Records after duplication removed (n =575)

Articles excluded using title and abstract (n=459)
Reasons:
- Articles on data coding, linkage or mining (n=12)
- Articles on other data assessment methods (n=12)
- Articles on electronic health records (n=15)
- Review articles (n=18)
- Quality of life/care studies (n=22)
- Articles on registry development and evaluation (n=39)
- Articles on clinical trial development and management (=63)
- Articles on database and technology infrastructure development (n=96)
- Articles on specific disease and treatments (n=182)

Records screened (n=619)

Records excluded if not available in English language and did not include human participants (n=183)

Title and abstract assessed for eligibility (n=575)

Full-text articles excluded (n=101)
Reasons:
- Data not from a clinical trial/research setting (n=38)
- Did not measure/report data quality through verifying original data with data entered to a CRF or electronic record (SDV) (n=63)

Full-text articles assessed for eligibility (n=116)

Studies included in critical appraisal (n=15)
<table>
<thead>
<tr>
<th>Author, year</th>
<th>SDV method</th>
<th>Source document to database</th>
<th>CRF to database</th>
<th>Source document to CRF</th>
<th>Frequency</th>
<th>Variables, n</th>
<th>Key data only</th>
<th>Data coded</th>
<th>Coding defined</th>
<th>Total error calculated</th>
<th>Outcomes/Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andrianopoulos et al. (2011)&lt;sup&gt;35&lt;/sup&gt;</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>F</td>
<td>Audit 1(a): 178(3); 1(b): 143(5); Audit 2:*&lt;sup&gt;(5)&lt;/sup&gt;</td>
<td>Audit 1: annually; Audit 2: N/A</td>
<td>Audit 1(a): 10; 1(b): 15</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Castelnuovo et al. (2012)&lt;sup&gt;14&lt;/sup&gt;</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>P</td>
<td>Pre: 100** Post: 10920**</td>
<td>Pre and post</td>
<td>4</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Christian et al. (1995)&lt;sup&gt;27&lt;/sup&gt;</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>P</td>
<td>1554 (86)</td>
<td>N/A</td>
<td>11</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Clarke et al. (2008)&lt;sup&gt;6&lt;/sup&gt;</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td>5(25)</td>
<td>N/A</td>
<td>20</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Cunningham et al. (2008)&lt;sup&gt;15&lt;/sup&gt;</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>P</td>
<td>942(94)</td>
<td>Annually</td>
<td>6</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>Duda et al. (2012)&lt;sup&gt;29&lt;/sup&gt;</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>P</td>
<td>184(86)</td>
<td>Re-audit of sites with major error</td>
<td>As many forms, a site could locate</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>Sample Size</td>
<td>Time Period</td>
<td>Data Fields</td>
<td>Error Rate</td>
<td>Monitoring Method</td>
<td>Finding/Conclusion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>------------</td>
<td>------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Herbert et al. (2004)</td>
<td>F</td>
<td>247(10)</td>
<td>6 months</td>
<td>315</td>
<td>x</td>
<td>√</td>
<td>Discrepancies occurred in &lt;5% of total data fields of the 99% of all audited charts</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horbar et al. (1995)</td>
<td>P</td>
<td>635(15)</td>
<td>N/A</td>
<td>10</td>
<td>x</td>
<td>√</td>
<td>Data keying errors reduced by introducing additional checks (logic, range and consistency).</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mealer et al. (2013)</td>
<td>P</td>
<td>32(30)</td>
<td>2 years</td>
<td>All primary and secondary endpoints</td>
<td>√</td>
<td>√</td>
<td>99% of data values were successfully monitored by remote SDV compared to traditional SDV.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mphatswe et al. (2012)</td>
<td>C</td>
<td>78(35)</td>
<td>Before, during and after</td>
<td>6</td>
<td>√</td>
<td>√</td>
<td>Completeness and accuracy of data improved from 1st to 3rd audit. Accuracy was determined by the degree of agreement.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nahm et al. (2008)</td>
<td>P</td>
<td>*(20-30)</td>
<td>Two-time points</td>
<td>11</td>
<td>√</td>
<td>×</td>
<td>Abstraction and transcription are most likely to introduce error. Error rates are dependent on the amount of data collected and the complexity of source documents.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tudur Smith et al (2012)</td>
<td>P</td>
<td>533**</td>
<td>N/A</td>
<td>All key variables</td>
<td>√</td>
<td>×</td>
<td>SDV is time-consuming, expensive, not necessarily error-free and identified random errors that had a minor impact. Central monitoring could be more effective than traditional SDV.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Van den Broeck et al (2007)</td>
<td>P</td>
<td>499(8)</td>
<td>Before, during and after (every 6 months)</td>
<td>6</td>
<td>√</td>
<td>x</td>
<td>Error rate &lt;1% set as a criterion for successful data. Difficulties included staff experience. Analysed data quality based on data flow stages.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>CRF</td>
<td>F</td>
<td>N/A</td>
<td>All items on CRF</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Incorrect data transfer was the main source of error. CRF design is important. Knowledge and experience of data management associated with % error.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-----</td>
<td>---</td>
<td>-----</td>
<td>------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>Accuracy of data elements ranged 79-99.5%. Data elements with suboptimal accuracy should be targeted for improvement.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vantongelen et al. (1989)34</td>
<td>✓</td>
<td>F</td>
<td>430 (100)</td>
<td>N/A</td>
<td>All items on CRF</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xian et al. (2011)32</td>
<td>✓</td>
<td>C</td>
<td>147(11)</td>
<td>N/A</td>
<td>35</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, year</td>
<td>Country</td>
<td>Research / Institutional publication</td>
<td>Publication Topic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------</td>
<td>----------------</td>
<td>--------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andrianopoulos et al. (2011)</td>
<td>Australia</td>
<td>Institutional</td>
<td>Cardiac surgery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Castelnuovo et al. (2012)</td>
<td>Africa</td>
<td>Institutional</td>
<td>HIV/AIDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Christian et al. (1995)</td>
<td>USA</td>
<td>Institutional</td>
<td>Cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clarke et al. (2008)</td>
<td>USA</td>
<td>Institutional</td>
<td>Cardiac surgery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cunningham et al. (2008)</td>
<td>New Zealand</td>
<td>Institutional</td>
<td>Cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duda et al. (2012)</td>
<td>South America</td>
<td>Institutional</td>
<td>HIV/AIDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Herbert et al. (2004)</td>
<td>USA</td>
<td>Institutional</td>
<td>Cardiac surgery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horbar et al. (1995)</td>
<td>USA</td>
<td>Research</td>
<td>Low birth weight</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mealer et al. (2013)</td>
<td>USA</td>
<td>Research</td>
<td>Acute respiratory syndrome and rare paediatric liver disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mphatswe et al. (2012)</td>
<td>Africa</td>
<td>Institutional</td>
<td>HIV/AIDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nahm et al. (2008)</td>
<td>USA</td>
<td>Institutional</td>
<td>Drug abuse treatment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tudur Smith et al (2012)</td>
<td>UK</td>
<td>Research</td>
<td>Cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vantongelen et al. (1989)</td>
<td>Europe</td>
<td>Research</td>
<td>Cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xian et al. (2011)</td>
<td>USA</td>
<td>Institutional</td>
<td>Stroke</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

HIV/AIDS: Human Immunodeficiency Virus/Acquired Immunodeficiency Syndrome; UK: United Kingdom; USA: United States of America
Research publication: Publications affiliated with academic research organisations
Institutional publication: Publications affiliated with only the pharmaceutical industry