

# Implementing error rate checks to improve the data quality in the Victorian Audit of Surgical Mortality

Andrew Chen, Jessele Vinluan\*, Claudia Retegan, Philip McCahy

Victorian Audit of Surgical Mortality, Royal Australasian College of Surgeons, East Melbourne, Victoria, Australia

## ARTICLE INFO

### Keywords:

Quality assurance  
Error rate  
Case record form  
Data field  
Visual inspection  
Source-to-database  
Electronic data capture  
Good clinical data management  
Good clinical practice

## ABSTRACT

**Objective:** The Victorian Audit of Surgical Mortality (VASM) is part of the Australian and New Zealand Audit of Surgical Mortality, aiming to identify factors associated with surgical mortality. A quality assurance method was utilised to rectify system issues through error rate checks of VASM data points. This method ensured data collected and entered in the audit database were accurate.

**Method:** Clinical data was collected using case record forms (CRFs), entered from paper forms by VASM staff or by the surgeon into an online interface. Closed audited cases from 1 January 2012 to 31 December 2016 were analysed. From 5528 closed cases, 485 (8.8%) were reviewed, containing a total of 1117 CRFs and 169,789 data fields. Each CRF entered was visually inspected in the database against the source document by an independent audit staff member. The error rate for each period was calculated and considered acceptable below 10 per 10,000 fields.

**Results:** Text errors made up the majority of data entry errors regardless of how the case was submitted.

**Conclusion:** Application of error rate checks is beneficial to maintain good clinical data. This activity improved and streamlined the data collection process to reduce errors associated with data entry. Once the entry system was stabilised, a reduction in error rate was observed showing potential for further improvements. We acknowledge that errors cannot be entirely eliminated and it is unrealistic. By investigating the rigour of the data management processes based on research guidelines, the findings can contribute to improve quality of clinical audits.

## 1. Introduction

The Victorian Audit of Surgical Mortality (VASM) is part of the Australian and New Zealand Audit of Surgical Mortality (ANZASM). The VASM monitors and identifies trends and outcomes to improve patient care [1]. Since inception the VASM utilised paper-based case record forms (CRF) and more recently implemented an electronic data capture (EDC) platform [2], known as the Fellows Interface, to collect clinical data from its stakeholders. To ensure the quality of surgical mortality data collected in the Bi-National Audit System (BAS) database are accurate, a quality assurance method has been implemented to rectify all system and manual entry errors. The VASM processes apply the operating principles of quality clinical registries [3], an adapted version of the International Clinical Harmonization Good Clinical Practice guideline [4], and clinical data management methodology [5] utilised in clinical trials.

Clinical data management methodology is integral in clinical research. Having a well-constructed database and CRF with

implementation of quality assurance tools such as validation and verification checks, clinical data review, staff training, maintenance and utilisation of standard operating procedures reflect good data management practices [4,6].

The most commonly used data collection tools are paper or electronic. Formerly, most health research organisations utilised paper-based data collection forms entered in a data repository and cross-checked for data verification [5,7–10]. In clinical trials, data quality is reviewed by sponsors to adhere to regulatory standards while in clinical registries, quality checks are driven by government consultants or by ensuring data integrity for future studies of clinical outcomes [4,9,11–14].

One of the challenges health researchers have in obtaining clean and reliable data is methodological checking for erroneous entries. Data processing errors are a common occurrence in health information databases and can often be attributed to human or system errors. Developments in rectifying data processing issues grew with the progression of information technology systems through automated system

\* Corresponding author. VASM - Royal Australasian College of Surgeons, GPO Box 2821, Melbourne, VIC 3001, Australia.

E-mail address: [Jessele.Vinluan@surgeons.org](mailto:Jessele.Vinluan@surgeons.org) (J. Vinluan).

validations. Studies investigated the benefits of electronic based systems and data reliability by scrutinising the methodology using double-data-entry or single-data-entry tests. Automated systems with built-in validation checks can provide a degree of utility; however, the types of errors relevant to data entry personnel, poor training, or poor CRF design were not identified [6,8,11,15,16].

An aspect of the clinical data management process adapted within the VASM process—and the focus of this paper—is the utilisation of error rate checks to visually inspect the entered data points. Errors can be minimised using quality checks to monitor and identify processing errors. The error rate check uses a source-to-database method commonly practiced in clinical trials and information technology. The error rate checking process involved the comparison of data points from the original source document to the data points entered into the system. This form of assessment allows for independent review of data integrity [5,7,9,16].

This paper aims to observe whether the introduction of the Fellows Interface, the EDC platform, reduces data errors over time based on the error rates and the impact of error rate checking as a quality assurance tool for data verification. We anticipate these findings will contribute to research in clinical audits by investigating the rigour of the data management processes according to research guidelines.

## 2. Methods

This study retrospectively analyses the VASM cases audited from 1 January 2012 to 31 December 2016. Over this period, the VASM collected data points using a mixture of paper and electronic CRFs. Data points were collected from several surgical specialties (Cardiothoracic, General, Ophthalmology, Oral and Maxillofacial, Orthopaedics, Otolaryngology, Paediatric, Plastic, Urology and Vascular), as well as Gynaecological cases. Three types of CRFs were utilised: the Surgical Case Form (SCF), First-Line Assessment (FLA) Form and Second-Line assessment (SLA) Form.

During the study period, the VASM audit process was triggered by surgical mortalities reported to the VASM office on a monthly basis by public and private Victorian healthcare services. Deaths were also self-reported by the treating surgeon or obtained from the Coroner's Court of Victoria and captured in BAS. For each surgical death entered in the system, a SCF was generated and sent to the treating surgeon to document their reflection of the patient's hospital admission and course to death. The SCF underwent a FLA by a surgeon-assessor. Where the first-line assessor requested a more detailed SLA, the patient hospital notes were obtained from the health service to be de-identified (removal of hospital, hospital staff and patient relatives' names, phone numbers, suburbs, signatures, and logos) before being sent to another surgeon-assessor for peer review. The audit loop was closed when feedback from the peer review was provided to the treating surgeon, forming the educational component of the audit.

The VASM initially utilised paper-based CRFs and data points were entered by audit staff into a customised Structured Query Language (SQL) server database, BAS. Data points manually entered were verified using Adobe's Optical Character Recognition (OCR) software, Intelligent Character Recognition recognising manuscript data with or without spaces in between Cardiff Teleform's Optical Mark Reading.

In 2011 the Fellows Interface, an EDC platform, was implemented enabling surgeons to enter data points directly onto an electronic SCF or FLA form. As part of the VASM's quality assurance, data entry checks were conducted on all data collection forms during the study period. Each CRF entered was visually inspected in the database against the source document by an independent audit staff member. This process was also applied to a scheduled error rate check on all processed forms. Initially, the aim was to conduct error rate checks on five per cent of the total surgical mortality cases which had been peer reviewed and closed. However, this sample for each error rate period was increased slightly to allow for more cases to be checked. The sample was selected

randomly to represent data points entered for that period. The sample was stratified based on the staff whose case it was allocated to (case owner) and the type of forms (either a SCF and FLA or a SCF, FLA and SLA).

The data errors were categorised as numeric, text or transcription. A numeric error was an incorrect entry of the digit or incorrect selection of a choice data field. A text error was an incorrect entry of a legible handwriting or printed text. Where entry was not a distinct numeric or text error such as illegible writing, it was considered a transcription error categorised as a discrepancy. In terms of electronic transcription, discrepancies were usually in the form of a typographical error, formatting issue, or misplaced text as a result of a scanned report copied into the database without being properly verified. The data types checked were free text, radio buttons (allowing for only one option to be selected), check boxes (allowing for multiple options to be selected) and selection lists.

During error rate checks, data fields with medical categorisation were also monitored. Read Codes were used to categorise the admission diagnosis, surgical diagnosis, cause of death, operative procedures, and deficiencies of care. This is a clinical decision tree that contains terms, synonyms, and abbreviations covering all aspects of patient care [17]. In an event of a data discrepancy, another independent checker from the audit office was consulted and the data error was re-categorised based on the context of the clinical information. Final verification of these cases was performed by the VASM Clinical Director with relevant clinical expertise to rectify the categorisation.

The results from the error rate checks were initially logged in a Microsoft Excel spreadsheet. From January 2015, findings were recorded in a Microsoft Access database developed to facilitate the recording of data points and generation of the report. The checker recorded the type of forms inspected, and whether the CRF data source was paper or electronic. The error log included the types of errors identified (e.g. a numeric or text error), and the initials of the case owner, data entry staff and data entry checker. The error rate for each period checked was calculated by identifying the number of errors multiplied by 10,000 then dividing this figure by the total number of data fields checked for each form [2]. An error rate below 10 per 10,000 fields was considered an acceptable level of error [5].

## 3. Results

A total of 5528 cases were closed during the study period, with a total of 12,057 CRFs received (Table 1). Of these, error rate checks were conducted on 8.8% (485/5528) cases. As per the audit protocol, all closed cases had a SCF and underwent a FLA. Of the cases selected for error rate checks, 32.0% (155/485) of these included a SLA review.

Of the 485 cases selected to be error rate checked, the total number of forms checked was 1117 with a combined total of 169,789 data fields as shown in Table 2. In total, 43.4% (485/1117) were SCF forms, 42.7% (477/1117) FLA forms, and 13.9% (155/1117) SLA forms (Fig. 1).

The type of errors identified over the study period were different based on the method of data submission, whether on a paper form or via the Fellows Interface.

### 3.1. Surgical case form (SCF)

The lowest total error rate recorded for SCFs were in 2016, with 7.0 errors per 10,000 fields, compared to the peak total error rate recorded in 2015 at 23.2 errors per 10,000 fields. Of these errors in 2015, 43 (16.9 errors per 10,000 fields) were on text fields and 16 (6.3 errors per 10,000 fields) on numeric fields.

The increased error rate in 2015 was the result of identifying a higher volume of text errors on electronic SCFs being submitted [13].

**Table 1**  
Distribution of Case Record Forms by form type and electronic/paper (n = 12,057 from 5528 cases).

| CRF Type |    | 2012        |              | 2013        |              | 2014        |              | 2015        |              | 2016        |              | Total       |              |
|----------|----|-------------|--------------|-------------|--------------|-------------|--------------|-------------|--------------|-------------|--------------|-------------|--------------|
|          |    | n           | (%)          | n           | (%)          | n           | (%)          | n           | (%)          | n           | (%)          | n           | (%)          |
| SCF      | E: | 302         | 30.8%        | 417         | 40.6%        | 474         | 47.9%        | 724         | 55.1%        | 711         | 58.3%        | <b>2628</b> | <b>47.5%</b> |
|          | P: | 677         | 69.2%        | 610         | 59.4%        | 515         | 52.1%        | 590         | 44.9%        | 508         | 41.7%        | <b>2900</b> | <b>52.5%</b> |
| FLA      | E: | 565         | 58.2%        | 672         | 65.6%        | 644         | 65.6%        | 869         | 66.1%        | 863         | 70.9%        | <b>3613</b> | <b>65.6%</b> |
|          | P: | 405         | 41.8%        | 353         | 34.4%        | 337         | 34.4%        | 445         | 33.9%        | 355         | 29.1%        | <b>1895</b> | <b>34.4%</b> |
| SLA      | E: | N/A         | –            | N/A         | –            | N/A         | –            | N/A         | –            | N/A         | –            | N/A         | –            |
|          | P: | 166         | 100%         | 185         | 100%         | 170         | 100%         | 248         | 100%         | 252         | 100%         | <b>1021</b> | <b>100%</b>  |
| Total    | E: | <b>867</b>  | <b>40.1%</b> | <b>1089</b> | <b>48.7%</b> | <b>1118</b> | <b>52.2%</b> | <b>1593</b> | <b>55.4%</b> | <b>1574</b> | <b>58.5%</b> | <b>6241</b> | <b>51.8%</b> |
|          | P: | <b>1248</b> | <b>59.0%</b> | <b>1148</b> | <b>51.3%</b> | <b>1022</b> | <b>47.8%</b> | <b>1283</b> | <b>44.6%</b> | <b>1115</b> | <b>41.5%</b> | <b>5816</b> | <b>48.2%</b> |

Note: n: number, CRF: case record form, SCF: surgical case form, FLA: first-line assessment, SLA: second-line assessment, E: electronic; P: paper. The significance of the bold numbers indicate the total numbers for each section.

### 3.2. First-line assessment (FLA) form

The total error rate for FLA forms fluctuated more than the SCFs, with 3.7 errors per 10,000 fields in 2012 and 35.8 errors per 10,000 fields in 2014. Of these errors in 2014, 12 (18.7 errors per 10,000 fields) were on text fields and 11 (17.1 errors per 10,000 fields) on numeric fields.

During the investigation, it was identified that one of the questions on the assessment form was restructured in 2014. When the restructure occurred, the question was updated in the BAS and on the paper form. However, for previous versions of the paper form that were already sent out, the question was presented differently on the electronic format of the BAS window. This resulted in the same error being picked up across multiple cases, causing the high error score for both the FLA and SLA forms in 2014. The number of errors decreased in 2015, as shown in Fig. 1, once staff members discussed and identified methodologies of verification to eliminate paper versus electronic field discrepancies and the BAS electronic system was also updated.

### 3.3. Second-line assessment (SLA) form

SLA forms saw the most fluctuation ranging from 7.1 errors per 10,000 fields in 2012 to 84.4 errors per 10,000 fields in 2014. Of these errors in 2014, 20 (64.9 errors per 10,000 fields) were on text fields and 6 (19.5 errors per 10,000 fields) on numeric fields. The spike in error rate can be seen up until 2014, before the rate dropped considerably in 2016. The SLA form contained a large amount of qualitative reports therefore there were a large amount of hand written information in submissions, some of it illegible. An example is provided in Fig. 2. The spike in error was largely due to lack of attention to detail on data entry forms and inadequate visual verification of the contents after transcripts were scanned using OCR.

### 3.4. Summary

When comparing the data submission type of all the CRFs checked, there were slightly more paper based forms compared to electronic form submissions, 56.8% (634/1117) paper versus 43.2% (483/1117) electronic.

In total, there were 308 errors identified during the audit study period, made up of 75.6% (233/308) text errors and only 24.4% (75/308) numeric errors. Text errors made up the majority of errors regardless of submission type, 78.2% (97/124) electronic and 73.9% (136/184) paper submissions.

One of the factors that influenced the error rate was the revision and system migration of the SCF form, leading to the paper and electronic versions being incongruent. Some examples include the introduction of

new data fields in the SCF, and moving the sub-questions relating to clinical management issues to a different section, as per an external evaluator's recommendations [18].

## 4. Discussion

Good clinical data quality entails the lowest possible number of errors and missing data points, to ensure the dataset is as complete and accurate as possible for analysis [5,6]. We acknowledge that there will always be room for error and it is unrealistic to expect complete elimination of human error [19]. The data entry errors found in this study were due to a combination of transcription errors [16] and misinterpretation of the information presented for data entry [8].

For example, errors in 2015 were mainly spelling mistakes that were not identified. The number of errors decreased in 2016 once staff members discussed and implemented methodologies of verification to reduce errors, as shown in Fig. 1.

Text errors made up the majority of data entry errors regardless of whether the case was received via paper or through the Fellows Interface. This concurs with previous research [2,7–9,11]. The average numeric and text error rates identified in a previous study were 14 and 976 per 10,000 respectively [7], which is slightly higher than the results obtained in this study. The reverification and revalidation of data points is a routine process done prior to analysis, as the text is utilised for qualitative publications such as case note review booklets [20].

Although there were two main categories of errors identified, they should be considered different based on how the data points were received. For example, text errors identified via paper form submissions arise from transcription issues due to illegible handwriting. Text errors identified from electronic forms arise from typographical errors not corrected during the data receipt stage and were classified under the same 'text errors' umbrella, although slightly different in the type of error they represent. Numeric errors were defined differently based on the type of form received. For example, it was easy to view and track the paper form errors as they were entered into the database incorrectly while electronic numeric errors were categorised as system errors.

The increase in error rate identified during the study period was partly due to the migration of BAS between different IT developers from late-2013 to early-2015. The enhancements caused some functionalities to become unstable. In 2015, an in-house developer was contracted to focus on the stabilisation of BAS which led to the decrease in error rate observed in 2016.

Consistency of data quality assessments are beneficial allowing improvements of the overall data quality [9]. Potential trends may be difficult to address if different data entry processes were utilised by the audit staff. Enforcing a standardised data entry process through training of staff members, including the independent checker should be part of a

**Table 2**  
Error rate per form type identified per 10,000 fields.

|                | 2012 |        |        | 2013 |        |        | 2014   |        |        | 2015   |        |        | 2016   |        |        | Total  |         |   |
|----------------|------|--------|--------|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|---------|---|
|                | E    | P      | T      | E    | P      | T      | E      | P      | T      | E      | P      | T      | E      | P      | T      | E      | P       | T |
| Fields/SCF     | 3    | 78     | 81     | 16   | 261    | 71     | 48     | 261    | 96     | 271    | 94     | 88     | 271    | 143    | 204    | 281    | 485     |   |
| SCF checked    | 783  | 20,358 | 21,141 | 4176 | 14,355 | 18,531 | 12,528 | 12,528 | 25,056 | 13,279 | 25,474 | 23,848 | 14,905 | 38,753 | 54,614 | 74,341 | 128,955 |   |
| Fields checked | 0    | 3.9    | 3.8    | 0    | 4.9    | 3.8    | 4.0    | 4.8    | 4.4    | 10.5   | 6.3    | 0      | 6.7    | 2.6    | 6.6    | 8.9    | 8.1     |   |
| Numeric        | 0    | 5.4    | 5.2    | 0    | 13.2   | 18.4   | 15.2   | 20.0   | 17.6   | 24.9   | 16.9   | 6.3    | 1.3    | 4.4    | 32.4   | 16.1   | 23.1    |   |
| Text           | 0    | 0      | 0      | 0    | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0       |   |
| Total          | 0    | 9.3    | 9.0    | 35.9 | 18.1   | 22.1   | 19.2   | 24.7   | 22.0   | 35.4   | 23.2   | 6.3    | 8.1    | 7.0    | 40.0   | 24.9   | 31.2    |   |
| Fields/FLA     | 14   | 67     | 81     | 37   | 67     | 63     | 58     | 67     | 96     | 62     | 94     | 99     | 62     | 143    | 279    | 198    | 477     |   |
| FLA checked    | 938  | 4489   | 5427   | 2479 | 1742   | 4221   | 3886   | 2546   | 6432   | 4402   | 5828   | 6138   | 2728   | 8886   | 17,843 | 12,931 | 30,774  |   |
| Fields checked | 0    | 4.5    | 3.7    | 0    | 0      | 0      | 18.0   | 15.7   | 17.1   | 2.3    | 3.4    | 0      | 0      | 0      | 10.5   | 9.2    | 9.5     |   |
| Numeric        | 0    | 0      | 0      | 0    | 28.7   | 11.9   | 20.6   | 15.7   | 18.7   | 4.5    | 20.6   | 8.2    | 33.0   | 15.8   | 17.6   | 56.0   | 28.3    |   |
| Text           | 0    | 0      | 0      | 0    | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0       |   |
| Total          | 0    | 4.5    | 3.7    | 0    | 28.7   | 11.9   | 38.6   | 31.4   | 35.8   | 6.8    | 24.0   | 8.2    | 33.0   | 15.8   | 28.1   | 65.2   | 37.7    |   |
| Fields/SLA     | 0    | 67     | 21     | 0    | 67     | 23     | 0      | 67     | 46     | 62     | 27     | 0      | 62     | 38     | 0      | 155    | 155     |   |
| SLA checked    | 0    | 1407   | 1407   | 0    | 1541   | 1541   | 0      | 3082   | 3082   | 0      | 1674   | 0      | 2356   | 2356   | 0      | 10,060 | 10,060  |   |
| Fields checked | 0    | 0      | 0      | 0    | 0      | 0      | 0      | 19.5   | 19.5   | 0      | 12.0   | 0      | 0      | 0      | 0      | 14.9   | 14.9    |   |
| Numeric        | 0    | 7.1    | 7.1    | 0    | 45.4   | 45.4   | 0      | 64.9   | 64.9   | 0      | 59.7   | 0      | 12.7   | 12.7   | 0      | 86.1   | 86.1    |   |
| Text           | 0    | 0      | 0      | 0    | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0       |   |
| Total          | 0    | 7.1    | 7.1    | 0    | 45.4   | 45.4   | 0      | 84.4   | 84.4   | 0      | 71.7   | 0      | 12.7   | 12.7   | 0      | 93.6   | 93.6    |   |

Note: ER: error rate, SCF: surgical case form, FLA: first-line assessment, SLA: second-line assessment, E: electronic; P: paper; T: total.  
The significance of the bold numbers indicate the total numbers for each section.  
Calculation: ER = ([ERROR]\*10,000)/([CHECKED]\*[FIELDS]).

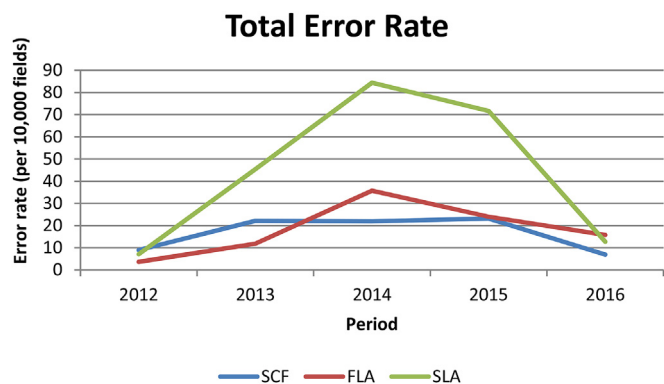


Fig. 1. Error rate by form type by year per 10,000 fields. Note: 10 per 10,000 errors as an acceptable level of error. SCF: surgical case form, FLA: first-line assessment, SLA: second-line assessment.

5407 - Patient severe cerebral palsy  
 Dx by another Sx at same hospital  
 Aug 2012 - underwent post op  
 decompression + Inst. Fusion.  
 Discharged home - small persistent  
 wound discharge

Represented to AdE sentic &  
 presumed chest wound & ongoing  
 minor distal wound weeping.  
 Mum wanted another Sx to offer  
 2nd opinion. I provided used  
 Sx agreed upon, in hope it would  
 improve sepsis. They were happy  
 chest was better. Mean CT  
 showed more C4-5 wound collect

Sx 6/10/12 - unremarkable  
 Present back to ICU Intubated  
 Bronchoscopy & washout in ICU 6/10/12  
 Extubated 7/10/12  
 Pt continues to improve slowly.  
 Answer notes: Best almost 10/10/12  
 ? PE ? Mucous Plug

Fig. 2. Text data entry example from a paper based CRF.

quality assurance activity in clinical research as education and awareness can impact on the outcome of the study. All professionals, including health clinicians, require training as new procedures, new equipment, new measurement and data collection processes are introduced to ensure protocol adherence and data integrity [15,21,22].

Conducting the error rate checks alone is not enough to result in any improvements to the data entry process. Additional validation and logic checks must be implemented methodically for monitoring the quality of the verified data points. The findings must be presented to all staff to avoid the recurring issues. This should generate technical discussions amongst staff members, allowing for comparisons to be made with previous reports to ensure the issues have been rectified successfully and any other potential concerns can be raised.

Identification of data errors are crucial factors of clinical databases, as it can disrupt the accuracy of any analysis being conducted. The presence of just a small percentage of data errors may lead to flawed

conclusions, misinterpretation of results or false positives, eventually needing correction [6]. This highlights the importance of acknowledging errors as they appear as a discrepancy in the dataset impacting data outcomes over time leading to inaccurate data analysis, potential bias and misrepresentation of results [9]. Error rate checks need to be considered as part of ongoing data quality maintenance to ensure data integrity of the research [6]. A well-structured data entry process was associated with improved accuracy of data entry. It is not unusual for data points to be submitted in a disorganised manner, which presents the possibility for data entry staff to become confused and enter it incorrectly [6]. The paper form versions needed to mirror the electronic system and vice versa [23].

Error rate check outcomes can assist in providing potential enhancements and developments of databases for data quality, and meet the needs for analysis and reporting of findings [9]. The assistance of EDC systems provide structure to the data entry process. For instance, the Fellows Interface's inbuilt validation rules can capture data errors and highlight inadvertently missed data entry to the clinician prior to submission. Audit staff are alerted to check their data entry on compulsory fields in BAS such as date of birth, date of admission to hospital, date of death and operation time and date. Monthly validation and logic checks were implemented on commonly found errors for specific data points. Thus, a prerequisite of the development and ongoing maintenance of these clinical systems should consider form design versions [15]. The utilisation of different form design versions allow for any revisions to the forms to be tracked, such as the addition or removal of clinical data points, user-friendliness and the provision of clear data entry instructions for clinicians as well as audit staff [6].

Despite increasing efforts in implementing electronic records, the risk for data entry errors to occur at the initial data entry stage can exist. Early detection and correction of errors will avoid misinformation from data point interpretations, so data quality assessments cannot be abolished [6,15,21].

Newly arisen issues with the migration towards electronic entry systems mean errors on the original source documents will become difficult to investigate and may result in different data management processes such as source document verification on site [7]. From the study findings, the implementation of the Fellows Interface experienced an initial rise in error rate due to the new processes, with the expected decrease of error rate observed over time [24]. The uptake of the Fellows Interface grew since its introduction in January 2012, jumping from 36.4% (334/917) to 68.2% (829/1216) in December 2016. The expected decrease of error rate throughout the study period has not yet fallen below the initial results in 2012. The downward trend in error rate once the BAS database was stabilised, along with the increased uptake of the online interface, demonstrated potential for further improvements.

#### 4.1. Limitations

One of the key limitations of this study was the subjectivity of the error rate checker. This could result in items being listed as an error rather than a difference such as BAS database errors due to system enhancements or grammatical errors. Examples for this include punctuation, spelling mistakes, and the tense used for sentences. For instance, the tense of the word from 'lead' to 'led' on an electronic SCF was marked as an error, but this could also be considered a difference depending on the error rate checker's judgement. 'Differences' are items that are debatable on whether the field should have been updated based on the information provided. It does not necessarily mean that the data entry was done in error, but could have been done differently.

Medical categorisation was not included in the scheduled error rate checks as this was independently verified by the Clinical Director with relevant expertise. Checking medical categorisation would require a separate activity and was beyond the scope of the study aims. Therefore any discrepancies in medical categorisation was recorded but not

further analysed or reported in the study findings.

The other limitation was the uniqueness of this method. This assurance activity was currently applied only in the Victorian region rather than on a national scale due to staff resources and cost feasibility.

The data for the FLA and SLA forms between October and December 2013 were excluded due to resourcing issues as no independent reviewer from the audit staff was available to conduct the error rate check.

## 5. Conclusion

Before electronic records become the standard, strategies need to be established to allow researchers to access the information directly and streamline the data collection process to, ideally, reduce errors associated with data entry. However, there are obstacles that need to be overcome; such as privacy issues, the amount of and type of data collected for the project [25], how the data will be utilised; also the method the researcher will use to access the data points needs to be considered based on research guidelines [7,23].

Providing sufficient training and ensuring staff members can improve the overall quality of the data points collected [8]. By carrying out error rate checks on data points being entered regularly, the integrity of the data points can be assessed and maintained appropriately. Error rate checking as a quality assurance activity has long term benefits – as a reliable source for publications. Data quality assurance tools are simple to implement and are able to provide valuable information on the accuracy of data entry [26].

The trend of error rate over the study period indicates a steady decrease after the online interface was implemented, which is expected to drop below the initial error rate in 2012.

It is easy to assume that the introduction of EDC platforms will automatically improve the data quality obtained, but this is not necessarily guaranteed. The key strength of conducting error rate checks is to identify any concerning trends and establish protocols to resolve these efficiently via soft checks, hard checks, logic checks, validation and source documentation checks.

Without quality data control, the number of erroneous publications and reports would dramatically increase and potentially mislead readers. All research organisations should monitor the quality of their data.

Since the completion of this study, the Fellows Interface has been nationally mandated. The continuation of this important quality assurance method, will allow for new types of errors that emerge to be rectified. It may be worthwhile repeating this study on a national level to identify data entry trends to discuss and appropriately address.

## Conflicts of interest

None declared.

## References

- [1] C. Beiles, C. Retegan, G. Maddern, Victorian Audit of Surgical Mortality is associated with improved clinical outcomes, *ANZ J. Surg.* 85 (2014) 803–807.
- [2] T. Jenkins, B. Wilson, R. Akers, et al., Evaluation of aTeleform-based data collection system: a multi-center obesity research casestudy, *Comput. Biol. Med.* 49 (1) (2014) 15–18.
- [3] Australian Commission on Safety and Quality in Health Care, Infrastructure and technical standards for Australian clinical quality registries sydney: Australian commission on safety and quality in health care, [2 Feb 2018]. Available from: <http://www.safetyandquality.gov.au/wp-content/uploads/2012/03/Infrastructure-and-Technical-Standards-for-Clinical-Quality-Registries-v1.0-May-2012.pdf>.
- [4] ICH Steering Committee, International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use: ICH Harmonized Tripartite Guideline: Guideline for Good Clinical Practice E6(R2), ICH Expert Working Group, Geneva, Switzerland, 1996 [2 Mar 2018]. Available from: <http://www.ich.org/products/guidelines/efficacy/efficacy-single/article/integrated-addendum-good-clinical-practice.html>.
- [5] The Society for Clinical Data Management, Good clinical data management practices (GCDMP) Belgium: the society for clinical data management, [2 Feb 2018]. Available from: <https://www.scdm.org/publications/gcdmp/>.

- [6] S. Matkar, A. Gangawane, An outline of data management in clinical research, *Int J Clin Trials* 4 (1) (2017) 1–6.
- [7] M.L. Nahm, C.F. Pieper, M.M. Cunningham, Quantifying Data Quality for Clinical Trials Using Electronic Data Capture, *Public Library of Science*, 2008 [cited 3 8]. e30491. Available from: [10.1371/journal.pone.0003049](https://doi.org/10.1371/journal.pone.0003049).
- [8] S.I. Goldberg, A. Niemierko, A. Turchin, Analysis of data errors in clinical research databases: AMIA annu symp proc, [19 May 2017]. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2656002/>.
- [9] R. Rostami, M. Nahm, C.F. Pieper, What can we learn from a decade of database audits? The Duke Clinical Research Institute experience, 1997–2006, *Clin. Trials* 6 (2) (2009) 141–150.
- [10] H. Li, S. Hawkl, K. Hanna, et al., Developing and Implementing a comprehensive clinical QA audit program, *Qual. Assur. J.* 11 (2007) 128–137.
- [11] M.K.H. Hong, H.H.I. Yao, J.S. Pedersen, et al., Error rates in a clinical data repository: lessons from the transition to electronic data transfer—a descriptive study, *BMJ Open* [Internet] 3 (5) (2013) e002406. Available from: [10.1136/bmjopen-2012-002406](https://doi.org/10.1136/bmjopen-2012-002406).
- [12] National Institute for Clinical Excellence, Principles for Best Practice in Clinical Audit United Kingdom, National Institute for Clinical Excellence, 2002 [2 Feb 2018]. Available from: <https://www.nice.org.uk/media/default/About/what-we-do/Into-practice/principles-for-best-practice-in-clinical-audit.pdf>.
- [13] C. Retegan, C. Russell, D. Harris, et al., Evaluating the value and impact of the victorian audit of surgical mortality, *ANZ J. Surg.* 83 (10) (2013) 724–728.
- [14] S.M. Evans, M. Bohensky, P.A. Cameron, et al., A survey of Australian clinical registries: can quality of care be measured? *Intern. Med. J.* 41 (1a) (2011) 42–48.
- [15] J. Neaton, A. Duchene, K. Svendsen, et al., An examination of the efficiency of some quality assurance methods commonly employed in clinical trials, *Stat. Med.* 9 (1–2) (1990) 115–124.
- [16] M.M. Wahi, D.V. Parks, R.C. Skeate, et al., Reducing errors from the electronic transcription of data collected on paper forms: a research data case study, *J. Am. Med. Inf. Assoc.* 15 (3) (2008) 386–389.
- [17] NHS Digital, Read codes: summary: NHS Digital, [31 Jul 2018]. Available from: <https://digital.nhs.uk/services/terminology-and-classifications/read-codes>.
- [18] Aspex Consulting, Evaluation of the Victorian Audit of Surgical Mortality (VASM) Melbourne, Royal Australasian College of Surgeons, 2015 [31 Jul 2018]. Available from: [https://www.surgeons.org/media/21916597/2015-07-28\\_rpt\\_vasm\\_external\\_evaluation\\_2015.pdf](https://www.surgeons.org/media/21916597/2015-07-28_rpt_vasm_external_evaluation_2015.pdf).
- [19] E.H. Bradley, L.A. Curry, S. Ramanadhan, et al., Research in action: using positive deviance to improve quality of health care, *Implement. Sci.* 4 (25) (2009).
- [20] Victorian Audit of Surgical Mortality, Case Note Review Booklet Melbourne, Royal Australasian College of Surgeons, 2018 [2 Mar 2018]. 10:[Available from: [https://www.surgeons.org/media/25626650/2018-02-06\\_rpt\\_vasm\\_cnr\\_b\\_v10\\_2018.pdf](https://www.surgeons.org/media/25626650/2018-02-06_rpt_vasm_cnr_b_v10_2018.pdf).
- [21] J.T. Mitchel, Y.J. Kim, J. Choi, et al., Evaluation of data entry errors and data changes to an electronic data capture clinical trial database, *Drug Inf. J.* 45 (4) (2011) 421–430.
- [22] M.R. Hamrell, B. Wagman, Standard operating procedures in clinical research: a beginner's guide, *Qual. Assur. J.* 5 (2) (2001) 93–97.
- [23] National Health and Medical Research Council, Australian Code for the Responsible Conduct of Research, National Health and Medical Research Council, Melbourne, 2007 [2 Feb 2018]. Available from: <https://www.nhmrc.gov.au/guidelines-publications/r39>.
- [24] G.M. Teasdale, Learning from Bristol: report of the public inquiry into children's heart surgery at Bristol Royal Infirmary 1984–1995, *Br J Neursurg* 16 (3) (2002) 211–216.
- [25] J.J. McNeil, S.M. Evans, N.P. Johnson, et al., Clinical-quality registries: their role in quality improvement, *MJA (Med. J. Aust.)* 192 (5) (2010) 244–245.
- [26] K. Walker, An Approach Finding the Right Tolerance Level for Clinical Data Acceptance Phoenix, Walker Consulting LLC, Arizona, 2005 [29 May 2018]. Available from: <https://www.lexjansen.com/pharmasug/2005/posters/po24.pdf>.

**Andrew Chen:** Mr Chen joined the Royal Australasian College of Surgeons (RACS) in 2011, and is currently a Project Officer for the Victorian Audit of Surgical Mortality (VASM). He has been involved in all aspects of the audit from improving data quality, data analysis for reports and publications, qualitative phone interviews with hospital stakeholders, training and mentorship of junior staff, as well as investigating new processes for electronic file sharing and reminders.

**Jessele Vinluan:** Miss Vinluan is a Senior Project Officer for the Victorian Audit of Surgical Mortality (VASM) at the Royal Australasian College of Surgeons (RACS) since 2007. Previously, she was involved in clinical audits, trials, and registries at the Baker Heart Research Institute's data management centre from 2002 to 2007. Her experience of recruiting patients into the registry studies led her to pursue a post-graduate training in mental health. She is a strong advocate for data management as a volunteer Executive Committee Member for the Australasian Health and Research Data Manager's Association (AHRDMA) since 2011; recently appointed as President in 2018.

**Claudia Retegan:** Ms Retegan commenced her research career at the Clinical Trials Research Unit, Auckland University as a Data Manager. She then became the Data Centre Manager at the Cardiovascular Disease Prevention Unit, Baker Heart Research Institute, Melbourne and focused on clinical trials, registries and audit programs. She joined the Royal Australasian College of Surgeons (RACS) where she became the Project Manager of the Victorian Audit of Surgical Mortality. She held several executive positions at the Australasian Health Research Data Managers Association and currently is a member of the RMIT Pharmaceutical Sciences Program Advisory Committee and the Victorian Healthcare Quality Association Committee.

**Philip McCahy:** Associate Professor McCahy is the Clinical Director of the Victorian Audit of Surgical Mortality as well as a practicing urological surgeon and adjunct senior lecturer at Monash University in Melbourne, Australia. He has numerous publications in peer reviewed journals and presentations at international scientific meetings. In addition he is heavily involved in developing urological services in Pacific Island nations.