

BMJ Open Quality Data quality audit of a clinical quality registry: a generic framework and case study of the Australian and New Zealand Hip Fracture Registry

Aidan Christopher Tan,¹ Elizabeth Armstrong,^{2,3} Jacqueline Close,^{2,4} Ian Andrew Harris^{5,6}

To cite: Tan AC, Armstrong E, Close J, *et al.* Data quality audit of a clinical quality registry: a generic framework and case study of the Australian and New Zealand Hip Fracture Registry. *BMJ Open Quality* 2019;**8**:e000490. doi:10.1136/bmjopen-2018-000490

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-000490>).

Received 26 July 2018
Revised 29 June 2019
Accepted 4 July 2019



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to
Aidan Christopher Tan;
aidan.tan@unsw.edu.au

ABSTRACT

Objectives The value of a clinical quality registry is contingent on the quality of its data. This study aims to pilot methodology for data quality audits of the Australian and New Zealand Hip Fracture Registry, a clinical quality registry of hip fracture clinical care and secondary fracture prevention.

Methods A data quality audit was performed by independently replicating the data collection and entry process for 163 randomly selected patient records from three contributing hospitals, and then comparing the replicated data set to the registry data set. Data agreement, as a proxy indicator of data accuracy, and data completeness were assessed.

Results An overall data agreement of 82.3% and overall data completeness of 95.6% were found, reflecting a moderate level of data accuracy and a very high level of data completeness. Half of all data disagreements were caused by information discrepancies, a quarter by missing discrepancies and a quarter by time, date and number discrepancies. Transcription discrepancies only accounted for 1 in every 50 data disagreements. The sources of inaccurate and incomplete data have been identified with the intention of implementing data quality improvement.

Conclusions Regular audits of data abstraction are necessary to improve data quality, assure data validity and reliability and guarantee the integrity and credibility of registry outputs. A generic framework and model for data quality audits of clinical quality registries is proposed, consisting of a three-step data abstraction audit, registry coverage audit and four-step data quality improvement process. Factors to consider for data abstraction audits include: central, remote or local implementation; single-stage or multistage random sampling; absolute, proportional, combination or alternative sample size calculation; data quality indicators; regular or ad hoc frequency; and qualitative assessment.

INTRODUCTION

Background

Hip fracture is a common and serious fall-related injury affecting older people. Almost all people who fracture their hip will be admitted to hospital and most will undergo a surgical procedure. An instrument to address the burden of hip fracture and instigate

large-scale change is the Australian and New Zealand Hip Fracture Registry (ANZHFR), a clinical quality registry of hip fracture clinical care and secondary fracture prevention. The ANZHFR receives core demographic, clinical and surgical information from before admission to after discharge for people aged 50 years and older admitted to participating hospitals.¹ The value of the ANZHFR is in coordinating and centralising binational collection of data that are not available in routine administrative data sets and are relevant to the management of patients with hip fracture. Consequently, the ANZHFR is able to measure current service provision, highlight the knowledge gap and maximise patient-centred outcomes after hip fracture by identifying areas for improvement.¹

A clinical quality registry is a systematic, standardised, structured and continuous collection of a prespecified minimum data set of health, process and outcomes data for people with particular health characteristics. By organising longitudinal, observational data from multiple participatory sites into a single central repository, clinical quality registries enable large-scale real-world register studies with greater statistical power, external validity and inferential reliability. The value of a clinical quality registry to inform care is contingent on the quality of the data it contains. The ANZHFR incorporates inbuilt data validation rules to identify overtly inaccurate data, and data completeness checks to identify incomplete data. This is assisted by a data dictionary and purpose-built data collection form to define and standardise data variables. Despite these data quality measures, some variables are not amenable to automated checks and the accuracy of the ANZHFR data has not been verified. High levels of data accuracy and data completeness will validate existing systems of preventing

and addressing suboptimal data, while low levels will indicate that existing systems are inadequate and novel improvement strategies are necessary.

Objectives

This study aims to pilot the methodology for a data quality audit of the ANZHFR. The primary aims are to (A) assess data agreement, as a proxy indicator for data accuracy, between the ANZHFR and a replicated data set collected by an auditor, and (B) appraise the data completeness of the ANZHFR. The secondary aim is to propose a generic framework and model for data quality audits in clinical quality registries.

METHODS

Study design

To assess the data accuracy and data completeness of the ANZHFR, a data quality audit was performed by independently replicating the data collection and entry process for 163 randomly selected patient records from three participating hospitals, and then comparing the replicated data set to the registry data set.

Setting

Between 1 January 2016 and 31 December 2016, thirty-four public hospitals contributed patient-level data to the ANZHFR. To minimise the confounding influence of inexperience with data collection, hospitals were eligible for inclusion in this study if they had participated in the ANZHFR for a minimum duration of 12 months. A subset of three public hospitals in one Australian state, representing 12% (3/25) of the 25 eligible hospitals, was pragmatically selected for their representativeness of the three most common data collection personnel: a single clinical nurse consultant, multiple geriatric trainees or multiple orthopaedic trainees.

Study size

Between 1 January 2016 and 31 December 2016, a total of 643 records were entered into the ANZHFR for patients admitted to the three selected hospitals: 185 from hospital A, 181 from hospital B and 277 from hospital C. Sample size was stratified by site and calculated using survey sample size methodology with a confidence level of 95% and an expected proportion of 0.95. This provided a CI of 0.05, with an upper bound of 1.00 and a lower bound of 0.90, and an SE of 0.03, with a relative SE of 2.69.² This achieved a sampling rate of 28.7% (53/185) for hospital A, 28.7% (52/181) for hospital B and 20.9% (58/277) for hospital C. The total sample size was 163 patient records, representing a total sampling rate of 25.4% (163/643).³

Participants

Records of patients eligible for selection were those admitted to hospital A, hospital B or hospital C between 1 January 2016 and 31 December 2016 inclusive. A second researcher, not directly involved with the ANZHFR data collection or entry, assigned the eligible records in the

registry data set an identification number.³ Patient records were stratified by site and selected using a random sequence generator. The second researcher provided the auditor with the identification number and the minimum patient information required to identify the medical record (ie, first name, last name, date of birth and medical record number). To protect privacy and confidentiality, patient identifying information was separated from patient clinical information, the abstracted data were only identified by the identification number and the abstracted data were secured in password-encrypted data spreadsheets only accessible to the auditor.

Variables

The data items collected by the ANZHFR are specified in the ANZHFR Data Dictionary v8.1.⁴ The replicated data set included all 53 independent data items under the first five categories: patient information, admission, assessment, treatment and discharge. Dependent data items derived from other data items were excluded, specifically age, length of stay in the operating hospital and length of stay in the hospital system. The two last categories, 30-day follow-up and 120-day follow-up, were excluded because they were not routinely reported and because abstraction would require contacting patients and be subject to recall bias.

Data sources

The registry data set was sourced from the ANZHFR. To avoid researcher bias, the auditor was blinded to the registry data set until the completion of data collection.

The replicated data set was produced by abstracting the data variables from the patients' paper and electronic medical records. To accurately replicate the registry data set, the auditor was trained by the most experienced data collector and the abstracted data were collected in strict accordance with the data dictionary (Data Dictionary v8.1) and entered in strict accordance with the data collection form (Patient Level Form v3). Data abstraction occurred at each hospital between May and August 2017.

Statistical methods

Statistical analysis was performed by matching and comparing the registry and replicated data set in Microsoft Excel 2016. To protect privacy and confidentiality, statistical analyses were stratified by data item, variable, category and hospital, but not by participant.

Data quality for registries comprised data accuracy, data completeness and registry coverage. Data accuracy refers to the correctness of the information recorded in the registry. Correctness refers to the closeness of the information recorded to the true value and is measured by the level of agreement between the registry data and the source data or verified third party data. In determining the data accuracy of the ANZHFR, the replicated data set was not considered the gold standard for data quality because it was impossible to exclude the possibility of the auditor abstracting inaccurate or incomplete data. Instead, data

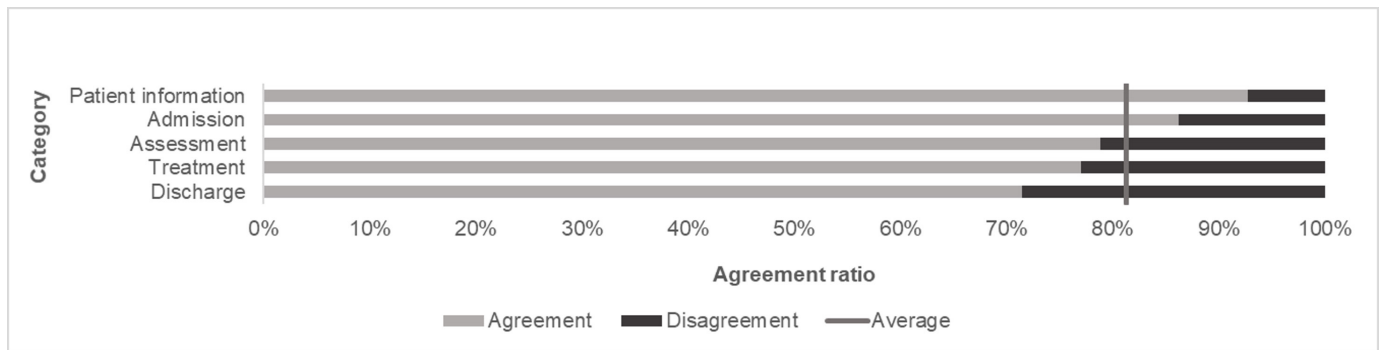


Figure 1 Agreement ratio by category.

agreement between the replicated and registry data sets was calculated as a proxy indicator for data accuracy. Missing data were coded and included in data analysis, such that for each variable there were a total of 163 possible matching entries. For categorical variables, representing most variables, agreement was defined as identical data in matching entries. For continuous variables, such as time, agreement is typically defined as similar data in matching entries, and data are considered to agree if it is within an acceptable range of values for which variation is not clinically relevant.³ In this study, a stricter definition of exact matching was applied to continuous variables because applying a reasonable window of error might overestimate data agreement and underestimate the potential for data quality improvement.

Data completeness refers to the extent to which the eligible data fields in a registration have a value recorded. Strict parameters were applied to data completeness, where both missing data and data coded as ‘not known’ were considered incomplete. While missing data, where the data field is left blank, suggest that the data were either unavailable or not sought, data coded as ‘not known’ suggest that the data were sought but not available or not found. Data coded as ‘not known’ were considered incomplete data because the implications for a registry are equivalent to missing data and to exclude these data might overestimate data completeness.

Benchmarks to gauge the level of data agreement and data completeness between the registry and replicated data sets for the ANZHFR were predetermined by a preceding literature review.^{5–10} Data agreement and data completeness are determined to be very high if $\geq 95\%$, high if $94\%–90\%$, high-moderate if $89\%–85\%$, moderate if $84\%–80\%$, low-moderate if $79\%–75\%$, low if $74\%–70\%$ and very low if $>70\%$.

Registry coverage refers to the extent to which the registry includes the target population. The percentage of registry coverage may be calculated as $100 - (\text{number of target population not included in the registry} \times 100 / \text{number of target population})$. The International Society of Arthroplasty Registries requires national registries to possess coverage of $\geq 80\%$.¹¹ While it was determined that 28.3% ($34/120$) of eligible public hospitals in Australia and New Zealand participated in the ANZHFR in 2016,¹

the coverage of participating patients as a proportion of eligible patients was outside the scope of this project.

To identify the frequency and causes of inaccurate data, data disagreements were stratified into six types of discrepancies: information, time, date, number, missing and transcription. Information discrepancies were defined as differences in the data recorded between the registry and replicated data sets in variables with coded non-numerical data. Time, date and number discrepancies were defined as differences in the data recorded between the registry and replicated data sets in variables with time data, date data or numerical string data, respectively. Missing discrepancies were defined as data missing or coded as ‘not known’ in either the replicated or registry data set, but where data were recorded in the corresponding data set. Transcription discrepancies were defined as one difference in the corresponding digits for numerical strings eight or fewer digits in length, and as one or two differences in the corresponding digits for numerical strings greater than eight digits in length.

Patient and public involvement

Patients and members of the public were not directly involved in the design or conduct of this study.

RESULTS

Data agreement

A total of 8639 data items were audited, representing 53 variables per record for 163 records. An overall data agreement, calculated as the total number of identically matching data items as a proportion of the total number of possible matching data items, of 82.3% ($7111/8639$) was found (figure 1). This represents a moderate level of data agreement.

When stratified by variable (figures 2 and 3), a median data agreement of 89.6% was found. Perfect data agreement was identified in 11.3% ($6/53$) of variables, and data agreement greater than 95% was identified in 43.4% ($23/53$) of variables. The two variables with the lowest data agreement, emergency department (ED) departure time (operating hospital) at 27.0% and time of surgery for hip fracture at 33.1% , were both time variables. When a ± 15 min window of error was applied to time data, data

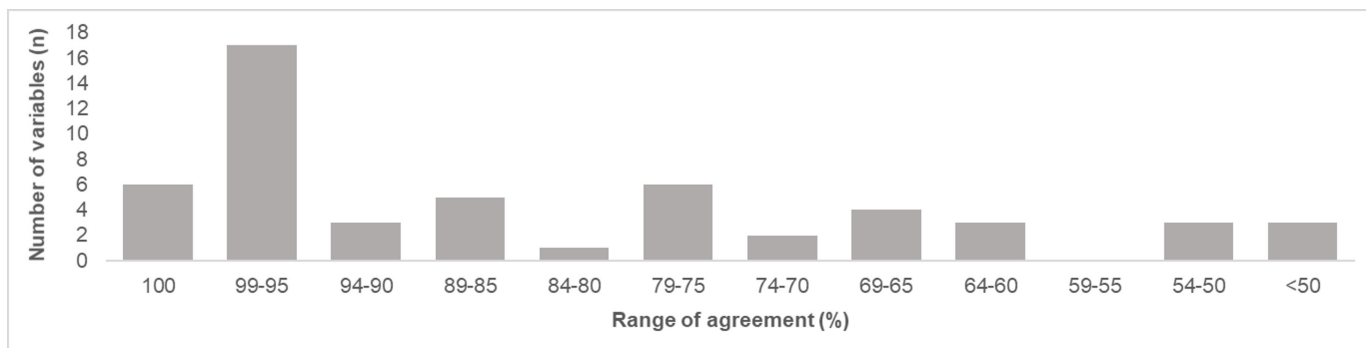


Figure 2 Number of variables by range of agreement.

agreement of these two variables increased to 44.2% and 39.3%, respectively, mean data agreement of time data increased from 63.9% to 73.7% and overall data agreement increased to 83.2%.

When stratified by site, similar overall levels of data agreement were found at 82.8% for hospital A, 79.3% for hospital B and 84.9% for hospital C.

When stratified by data disagreement type (figures 4 and 5), almost half, 46.2% (706/1528), of all data disagreements were caused by information discrepancies. Approximately a quarter, 27.3% (417/1528), of all disagreements were caused by time, date and number discrepancies: 12.8% (196/1528), 10.7% (163/1528) and 3.8% (58/1528), respectively. Approximately a quarter, 24.5% (374/1528), of all disagreements were caused by missing discrepancies. While transcription errors were among the most frequently cited causes of disagreeing data,⁵ only 1 in every 50 disagreements, 2.0% (31/1528), were caused by transcription discrepancies.

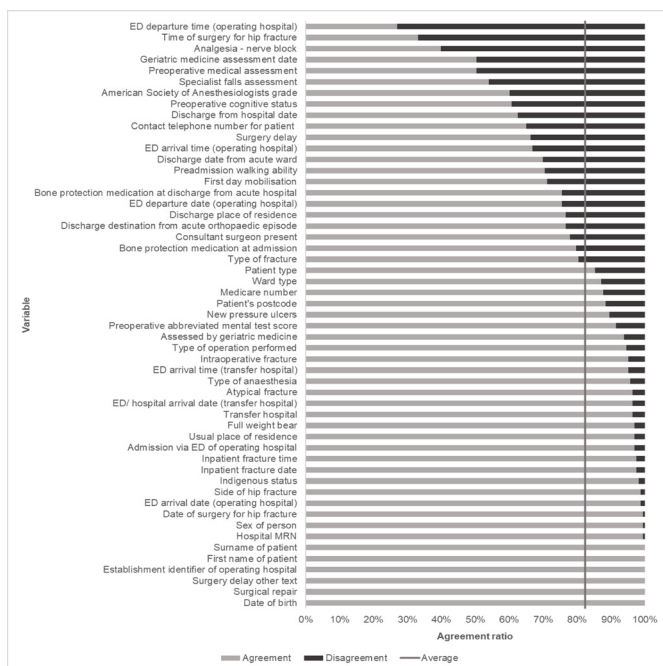


Figure 3 Agreement ratio by variable. ED, emergency department.

Data completeness

An overall data completeness, calculated as the total number of complete data items as a proportion of the total number of eligible data items, of 95.6% (8258/8639) for the registry data set, and 95.6% (8257/8639) for the replicated data set was found. This represents a very high level of data completeness. When excluding one variable, preoperative abbreviated mental test score, data completeness of the registry and replicated data set increased to 97.3% (8404/8639) and 97.3% (8405/8639), respectively, and retained equivalence.

When stratified by variables, a median data completeness of 100% for the replicated data set and 99.4% for the registry data set was found. Perfect data completeness was identified in 52.8% (28/53) of variables in the replicated data set and 47.2% of variables in the registry data set. Variables derived from paper medical records were associated with lower levels of data completeness and more missing discrepancies than variables derived from electronic medical records. The two variables with the lowest data completeness in the replicated data set were surgery delay and preoperative cognitive status, while the two variables with the lowest data completeness in the registry data set were discharge place of residence at 74.2% (121/163) and discharge from hospital date at 77.3% (126/163).

DISCUSSION

Using strict criteria, a moderate level of data agreement and a very high level of data completeness were found in the ANZHFR. When a ± 15 min window of error was applied to time data, overall level of data agreement increased by 1%. While discrepancies between the registry and replicated data set may represent an error in either data set, they are more likely to represent inaccurate or incomplete data in the registry data set given the strict parameters of exact matching and systematic approach applied to data abstraction. Using agreement as a proxy indicator of accuracy, a moderate level of data accuracy was inferred.

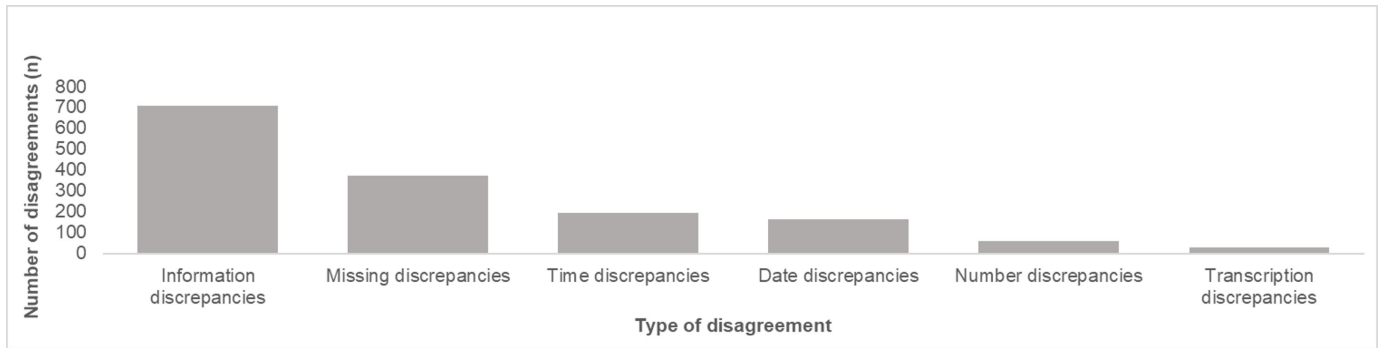


Figure 4 Number of disagreements by type of disagreement.

Data agreement

By stratifying data disagreements by type of discrepancy, sources of inaccurate data were identified. Information discrepancies are the product of either inadequately defined or inappropriately interpreted definitions of data. Vague or imprecise definitions are susceptible to misinterpretation. For example, one option for the variable of first day mobilisation was ‘patient out of bed and given opportunity to start mobilising day one post-surgery’. If a patient was provided with the opportunity to mobilise but did not mobilise, some data collectors considered this variable satisfied as the opportunity was provided, whereas other data collectors did not consider this variable satisfied as the patient did not mobilise. This may be addressed by using strict and unambiguous definitions of data, tightening and specifying the parameters of the data dictionary, incorporating the definitions of data into the data collection form, reviewing the definitions of data as part of a data quality audit, and training data collectors in data extraction, interpretation

and submission, particularly in variables with high data disagreement. The correlation between the chronological order of data generation, and data agreement and completeness may be explained by the nature of the variables within each category. Earlier categories include a higher proportion of administrative and other routinely collected and readily available data. However, later categories rely more on unique data that are documented in the medical record in a less systematic and standardised manner. This complicates data abstraction, and by extension data agreement, because the auditor would not have been involved in patient management and may find it difficult to locate all the required information from the patient medical record.

Time, date and number discrepancies are primarily the product of data collectors deriving data from divergent sources within a medical record or a non-documented source such as a clinician or the patient. The two variables with the lowest data agreement were time variables, ED departure time (operating hospital) and time of surgery for hip fracture, although data disagreement was demonstrated to be amplified by the treatment of time as a categorical variable in the context of multiple data sources with minor variations. The presence of multiple or non-reproducible data sources introduces potential variation by placing the responsibility for determining the correct data source on the data collector. This may be addressed by developing a detailed hierarchy of source documentation in the data dictionary, including the location of the primary data source at each site, with greater weight given to confirmation by objective evidence and relevant experts. This may also be addressed by directly exporting routinely collected time, date and number data from administrative databases into registry databases.

Missing discrepancies are the product of a divergence between the registry and the replicated data collection in identifying the relevant information. These disagreements may be addressed through the previously proposed hierarchy of source documentation in the data dictionary. Transcription discrepancies, a small cause of error in this registry, are a product of human error and may be addressed by increasing the awareness and attention of data collectors to transcription errors. These

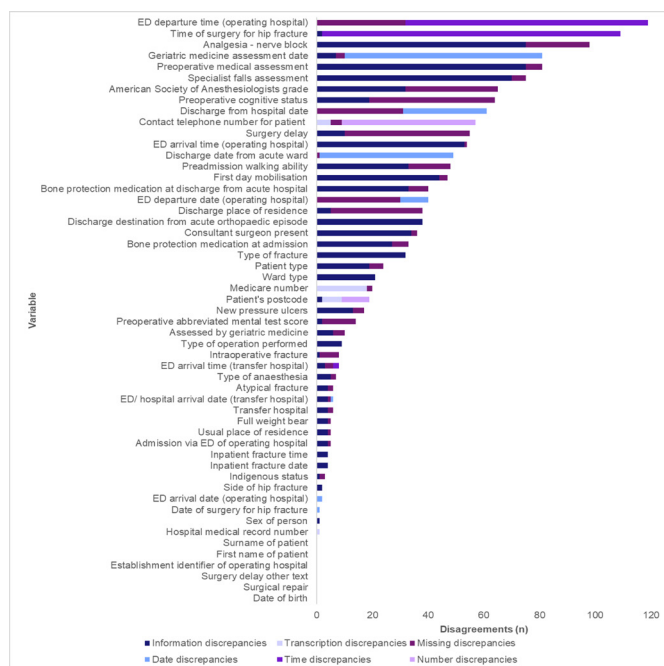


Figure 5 Number and type of disagreements by variable. ED, emergency department.

Table 1 Recommendations to address data disagreement

Data discrepancies	Recommendations
Information discrepancies	A. Use strict and unambiguous definitions of data. B. Tighten and specify the parameters of the data dictionary. C. Incorporate the definitions of data into the data collection form. D. Review the definitions of data as part of a data quality audit. E. Regular training sessions in data extraction, interpretation and submission.
Time, date and number discrepancies	A. Develop a detailed hierarchy of source documentation in the data dictionary. B. Directly export routinely collected time, date and number data from administrative databases into registry databases.
Missing discrepancies	A. Develop a detailed hierarchy of source documentation in the data dictionary.
Transcription discrepancies	A. Increase the awareness and attention of data collectors to transcription errors.

recommendations on addressing data disagreement are tabulated in [table 1](#).

While the ANZHFR uses different data collection personnel, data collection method is consistent across sites, in line with the current recommendations for clinical registries. Hospital C, using a single clinical nurse consultant data collection model, had the highest level of data agreement. While it is outside the scope of this study to draw conclusions as to whether this is the most accurate data collection model, the similar levels of data agreement between sites may suggest that different data collection personnel do not impact agreement.

Data completeness

The very high level of data completeness is likely secondary to the strict automated and manual data completeness checks implemented by the ANZHFR. Interestingly, despite the almost identical overall data completeness between the registry and replicated data sets, there was limited correlation and overlap between the incomplete data items in the registry data set and the incomplete data items in the replicated data set.

A contributing cause of data incompleteness and missing discrepancies is the format and presentation of the medical record. A disorganised arrangement of medical records or unclear handwriting may complicate data abstraction. Variables derived from paper notes written by clinical staff were associated with lower levels of data completeness and more missing discrepancies than typed administrative records. This overlaps with the type of data, as demographic data were associated with higher levels of data completeness compared with clinical data that required interpretation. Another contributing cause of missing discrepancies is an inability to access data sources because either the medical record is incomplete during data collection, the data variable is not documented or the information is sourced from the patient or family.⁶ For the two variables with the lowest data completeness in the replicated data set, almost one-third of data items were known by the data collectors for the registry data set but documentation in the medical record was not found by the auditor for the replicated data set. Correspondingly, the two variables with the lowest data

completeness in the registry data set may be explained by the shift in the involvement of the data collectors in patient care as the patient transitioned from acute care to rehabilitation. It is hoped that the adoption of electronic medical records in hospitals, with discrete data fields and inbuilt validation checks, will mitigate data entry errors, incomplete data and missing discrepancies. In the interim, extra effort in ensuring accuracy and completeness is necessary through structured training and ongoing education sessions, central and local data oversight, direct and regular data transfer, functional information technology infrastructure, automated range, consistency and completeness checks, and manual duplicate, visual, calculation and reliability checks.^{6 12}

Generic framework and model

Three broad origins of incorrect or incomplete data were identified: the data source; data parameters; and data collection, including both human and computer data collection and entry. A generic framework (online supplementary table A1) for data quality audits in clinical quality registries is proposed.

Limitations

This study assessed the agreement between the registry and replicated data sets but could not verify the accuracy of the data sources. Although the data documented in the electronic and paper medical records are the most reliable source available, they may have contained inaccurate data, or information may have been sourced from the patient or family at the time of admission and not documented therefore unavailable to retrospective review.⁵ Additionally, while the three selected hospitals represent the three most common data collection personnel, they were all public hospitals in a metropolitan area, and consequently the findings may not be generalisable to private hospitals or hospitals in regional or rural areas.

CONCLUSION

In the first data quality audit of the ANZHFR, a moderate level of data accuracy and a very high level of data completeness were found. Where possible, sources of

inaccurate and incomplete data have been identified with the intention of implementing data quality improvements. As the means and methods of data collection and entry undergo revision, data accuracy and data completeness are anticipated to increase over time. Regular audits of data abstraction are necessary to improve data quality, assure data validity and reliability, and guarantee the integrity and credibility of registry outputs. Through these efforts, registries can improve quality of care by evaluating clinical practice using reliable data.

Author affiliations

¹Faculty of Medicine, University of New South Wales, Sydney, New South Wales, Australia

²Falls, Balance and Injury Research Centre, Neuroscience Research Australia, Sydney, New South Wales, Australia

³School of Public Health and Community Medicine, University of New South Wales, Sydney, New South Wales, Australia

⁴Prince of Wales Clinical School, University of New South Wales, Sydney, New South Wales, Australia

⁵South Western Sydney Clinical School, University of New South Wales, Sydney, New South Wales, Australia

⁶Whitlam Orthopaedic Research Centre, Ingham Institute for Applied Medical Research, Sydney, New South Wales, Australia

Acknowledgements Associate Professor Boaz Shulruf and Dr Wei Xuan provided statistical support. Mr Stewart Fleming provided technical assistance.

Contributors Conceptualisation: IAH, EA and ACT. Methodology: IAH, EA and ACT. Formal analysis: ACT. Investigation: ACT. Data curation: ACT. Writing—original draft preparation: ACT. Writing—review and editing: ACT, IAH, EA and JC. Supervision: IAH.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests IAH and JC are co-chairs of the ANZHFR. EA is the Australian registry manager of the ANZHFR.

Patient consent for publication Not required.

Ethics approval Ethics approval was provided by the NSW Population and Health Services Research Ethics Committee (HREC/14/CIPHS/51).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No additional data are available.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

1. Australian and New Zealand hip fracture registry. ANZHFR bi-national annual report for hip fracture care, 2017. Available: <http://anzhfr.org/wp-content/uploads/2017/08/ANZHFR-Annual-Report-2017.pdf>
2. National Statistical Service. Sample size calculator: Australian Bureau of statistics, 2017. Available: <http://www.nss.gov.au/nss/home.nsf/pages/Sample+size+calculator>
3. Seagrave KG, Naylor J, Armstrong E, *et al*. Data quality audit of the arthroplasty clinical outcomes registry NSW. *BMC Health Serv Res* 2014;14:512.
4. Australian & New Zealand Hip Fracture Registry. Data Dictionary 8.1, 2013. Available: http://anzhfr.org/wp-content/uploads/2015/04/ANZHFR_DataDictionary_v8.1_December2013.pdf
5. Arts DGT, De Keizer NF, Scheffer G-J. Defining and improving data quality in medical registries: a literature review, case study, and generic framework. *J Am Med Inform Assoc* 2002;9:600–11.
6. Herbert MA, Prince SL, Williams JL, *et al*. Are unaudited records from an outcomes registry database accurate? *Ann Thorac Surg* 2004;77:1960–4.
7. van der Meulen JHP, Jacob M, Copley L. Assessing the quality of the data in a transplant Registry: the European liver transplant registry. *Transplantation* 2003;75:2164–7.
8. Kuwatsuka Y. Quality control and assurance in hematopoietic stem cell transplantation data registries in Japan and other countries. *Int J Hematol* 2016;103:20–4.
9. Moussa I, Hermann A, Messenger JC, *et al*. The NCDR CathPCI registry: a US national perspective on care and outcomes for percutaneous coronary intervention. *Heart* 2013;99:297–303.
10. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
11. International Society of arthroplasty registries. International Society of arthroplasty registries, 2018. Available: <http://www.isarhome.org/>
12. Australian Commission on Safety and Quality in Health Care. *Framework for Australian clinical quality registries*. Sydney: ACSQHC, 2014: 23–4.