Validating a transnational fracture treatment registry using a standardized method

Jasper Frese (jasper.frese@yahoo.de)
Universitatsklinikum Schleswig-Holstein Klinik fur Orthopadie und Unfallchirurgie
https://orcid.org/0000-0003-1062-2711

Annalice Gode
Universitatsklinikum Schleswig-Holstein Klinik fur Orthopadie und Unfallchirurgie

Gerhard Heinrichs
Universitatsklinikum Schleswig-Holstein Klinik fur Orthopadie und Unfallchirurgie

Armin Will
Universitatsklinikum Schleswig Holstein, Stabsstelle Informationstechnologie

Arndt-Peter Schulz
BG Klinikum Hamburg

Research article

Keywords: Data validation, Registry, quality assessment, scoring, data quality

Posted Date: November 18th, 2019

DOI: https://doi.org/10.21203/rs.2.11475/v4

Version of Record: A version of this preprint was published on December 18th, 2019. See the published version at https://doi.org/10.1186/s12874-019-0862-1.

Validating a transnational fracture treatment registry using a standardized method

Frese, Jasper; UKSH Campus Lübeck, Orthopaedics and Traumatology (corresponding author)
E-Mail: jasper.frese@yahoo.de
Gode, Annalice; UKSH Campus Lübeck, Orthopaedics and Traumatology
E-Mail: annalice.gode@uksh.de
Heinrichs, Gerhard; UKSH Campus Lübeck, Orthopaedics and Traumatology
E-Mail: gerhard.heinrichs@uksh.de
Will, Armin; UKSH Campus Lübeck, Stabsstelle Informationstechnologie
E-Mail: armin.will@uksh.de
Schulz, Arndt- Peter; Klinik für Orthopädie und Unfallchirurgie, BG Klinikum Hamburg
E-Mail: schulz@biomechatronics.de

Abstract

Aim

Subsequent to a three-month pilot phase, recruiting patients for the newly established BFCC (Baltic Fracture Competence Centre) transnational fracture registry, a validation of the data quality needed to be carried out, applying a standardized method.

Method

During the literature research, the method of “adaptive monitoring” fulfilled the requirements of the registry and was applied. It consisted of a three-step audit process; firstly, scoring of the

overall data quality, followed by source data verification of a sample size, relative to the scoring result, and finally, feedback to the registry on measures to improve data quality. Statistical methods for scoring of data quality and visualisation of discrepancies between registry data and source data were developed and applied.

Results

Initially, the data quality of the registry scored as medium. During source data verification, missing items in the registry, causing medium data quality, turned out to be absent in the source as well. A subsequent adaptation of the score evaluated the registry's data quality as good. It was suggested to add variables to some items in order to improve the accuracy of the registry.

Discussion

The application of the method of adaptive monitoring has only been published by Jacke et al., with a similar improvement of the scoring result following the audit process. Displaying data from the registry in graphs helped to find missing items and discover issues with data formats. Graphically comparing the degree of agreement between the registry and source data allowed to discover systematic faults.

Conclusions

The method of adaptive monitoring gives a substantiated guideline for systematically evaluating and monitoring a registry's data quality and is currently second to none. The resulting transparency of the registry's data quality could be helpful in annual reports, as published by most major registries. As the method has been rarely applied, further successive applications in established registries would be desirable.

Key words: Data validation, Registry, quality assessment, scoring, data quality

Background

From November 2017 until February 2018, a transnational fracture registry with a complication module was piloted at the Lübeck University Hospital (UKSH Campus Lübeck, Germany), within the framework of the European Union (EU) funded Baltic Fracture Competence Centre (BFCC) Project[22]. A novel classification for complications during fracture treatment was applied and a follow-up letter was shipped out to all registered patients 6 months after treatment. Subsequent to the registration and follow-up phase, an assessment of the registry's data quality needed to be carried out. Assessment and monitoring of a registry's data quality are crucial to make it a reliable tool to be used[20] and to reinforce trust in the research conducted[25]. As a transnational registry, data quality assessment needed to follow a standardized procedure to allow comparability between data entering centres. Nonnemacher et al.[19] have developed the method of "adaptive monitoring" that fulfilled the requirements by the registry. The following article documents its application in detail and proposes a statistical method for finding systematic faults during source data verification.

Materials

Setting and Data Capture

During the pilot phase for the BFCC fracture registry at the Lübeck University Hospital from mid-November 2017 to mid-February 2018, physicians enrolled 238 patients with fresh fractures, meaning less than a week old and not treated prior to admission, of the extremities and pelvis (of any type). Exclusively adult patients that had understood the nature of the study and consented to be entered into the registry were included. A distinguishing feature was the registration of adverse events, here synonymous to complications. Data capture was carried

out manually, using registration forms that were later entered into the database by study nurses. The registry was web-based and had its own information technology (IT) infrastructure, providing high security standards by using browser client security certificates, pseudonymization of patients, and separate storage of patient and medical data by a trusted third party. The software used was Centraxx by Kairos[14]. Furthermore, six months post-treatment, patients were contacted via mail and asked, in a pseudonymized follow-up questionnaire, whether they were satisfied with their treatment, had experienced any complications, and if so, how defacing they were. The questionnaire was sent back to the hospital by postage-paid mail or as a scan to an e-mail address.

Finding a method for data quality assessment

Initially, an extensive literature research was conducted, using PubMed[3], MEDLINE, Google Scholar[7] and Medscape[24], for publications containing the following keywords: registry, data validation, data quality, standardized, and method. Furthermore, reports of established registries, including the in-hospital Stavanger fracture and dislocation registry[18], Vascular Registry of Denmark[16], prospective registry for surgical complications in the Surgical Department of St. Elisabeth's Hospital in Tilburg, Netherlands[23], Swedish Fracture Registry[13], Swedish Cancer Registry[9], Australian Orthopaedic Association National Joint Replacement Registry[8], Danish Cancer Registry[6], German Multiple Sclerosis Registry[5], and Cancer Registry of Norway[15] were studied regarding their methods, parameters investigated and results of data quality validation.

Established registries used varying methods to evaluate their data quality, ranging from reproducibility of entered data sheets on cases[16], completeness and correctness of registered procedures and diagnoses[8, 13, 18] up to plausibility of entered datasets[5]. No two registries had

entirely matching approaches to data validation and no standardized systematic using transparent and reproducible methods could be identified.

On a German platform for techniques and methodology in medical research[1], the method of adaptive monitoring developed by Nonnemacher et al.[19] was found, suggesting a standardized approach to data validation called adaptive monitoring. The three-step audit process consisted of firstly scoring of the overall data quality, followed by source data verification (SDV) of a sample size relative to the scoring result and finally, a feedback to the registry on measures to improve data quality. The application of this method to the BFCC registry was documented in detail.

Methods

Scoring data quality and conducting a source data verification

To determine the score for the evaluation of the data quality, all registered patients were included in the calculation. Three levels of data quality were investigated: organization, integrity, and correctness. Each level had indicators assigned to it, which were evaluated according to a predefined threshold for sufficient data quality: on the level of organization, the qualification of data entering personnel was evaluated, and on the level of integrity, the optional data elements height and weight were searched for missing entries. Here, the body mass index (BMI) was calculated and the proportion of non-calculable elements identified. Furthermore, value distributions of mandatory elements were investigated. Here, the length of stay was calculated and analysed using the graphical method *geom_density* for the package ggplot2 in R Statistics[2], with the length of stay in days on the *x*-axis and the density of patients on the *y*- axis. On the level of correctness, measurable elements of the inclusion criteria, age of the fracture and patient age at the time of inclusion in the study were checked. For investigating

the fracture age, the same method as for analysing the length of stay was applied. Indicators had specific thresholds, defining whether the data quality was sufficient. When passing acceptance levels, a factor of 1 was assigned, and when failing, a factor of 0, which was then multiplied by their specific predefined weight. Items, indicators, thresholds and weights are summarized in Table 1[19].

Items were analysed using the statistical software R Statistics[2] (Version 3.5.1) and RStudio (both R Consortium, Boston, MA, USA). Subsequent to investigating all items, a score was calculated by dividing the sum of the results of each indicator ($\boldsymbol{IW}$ = Individual Weights) by the sum of their specific weights ($\boldsymbol{SW}$ = Sum of Weights) multiplied by 100; summarized as formula 1:

$$\boldsymbol{Score} = \frac{\boldsymbol{IW}}{\boldsymbol{SW}} \times \mathbf{100}$$

Scoring results were stratified to evaluate the data quality from very poor to very good, and in addition, matched with a recommended factor δ (delta), which was used to calculate the number of patients, upon which SDV should be conducted. Table 2[19] summarizes delta values in relation to score results and data quality.

The unadjusted case number $\boldsymbol{n_0}$ (independent of the cohort size) needed to be calculated as a foundation for the adjusted case number $\boldsymbol{n}$ (the sample size taken from the registry for SDV), which was relative to the total number of patients in the registry.

Formula 2 was used to calculate the $\boldsymbol{n_0}$ for SDV:

$$\boldsymbol{n_0} = \frac{\mathbf{p(1-p)}}{\boldsymbol{\delta^2}} \times \boldsymbol{z^2_{1-\alpha/2}}$$

For the first SDV, Nonnemacher et al. recommended a $\mathbf{p}$-value$= \mathbf{0.05}$. For the quantile of the standard normal distribution $\boldsymbol{z^2_{1-\alpha/2}}$ for a first-order error of $\alpha = 0.05$, Nonnemacher et al. recommended $\boldsymbol{z^2_{1-\alpha/2}} = \mathbf{1.96}$.

Finally, $\boldsymbol{n}$, defining the sample size taken from the registry for SDV, related to $\boldsymbol{n_0}$ and to the total number of patients registered, defined as $\boldsymbol{N}$, was calculated using formula 3:

$$\boldsymbol{n} = \frac{\boldsymbol{n_0} \cdot \boldsymbol{N}}{\boldsymbol{n_0} + \boldsymbol{N}}$$

Conducting a source data verification

Within the statistical software R Statistics[2], the package ggplot2 and its *geom_count* and *geom_density* functions were the method of choice to investigate systematic faults while conducting the SDV. The SDV was shown on the *x*-axis, registry data on the *y*-axis, and 'same' on the x- axis indicating a match between the registry and the source data. The parameter "same" was introduced as a constant for each item investigated in order to make the method applied more comprehensible. The legend abbreviation 'prop' indicated the proportion of agreement between the registry and the source data. If a circle equalled a prop of 1.0, the match between the registry and source was 100%. To facilitate understanding the results of the SDV, an arbitrary example displaying 'The perception of personnel by patients in comparison to the actual function of personnel' is given in Figure 1. The following conclusions can be drawn from this example: All (100%) of the nursing personnel was also perceived as nursing personnel (large green circle). Half of the doctors were identified as students (left blue circle), and 25% of the students were actually doctors (small red circle). During SDV, deviations from source to registry of double-digit percentiles were further analysed using this graphical method.

The SDV focused on 12 selected items from the registry, ranging from general medical information, administrative entries to fracture-specific data, such as admission date, discharge date, date of treatment, height and weight, employment status, fracture side, number of comorbidities, main diagnosis according to International Statistical Classification of Diseases and Related Health Problems German Modification (ICD-10 GM), fracture date, occurrence of

a complication, type of fixation, and type of reduction. It was carried out on each sampled patient and analysed using R Statistics[2].

Results

Scoring data quality

Items for scoring the data quality were analysed, individual weights calculated, and the results summarized in Table 3. On the level of organisation, all personnel at the study centre had been diligently trained; the result was 100% and the individual weight of 2 calculated. The length of stay, shown in Figure 2, hinted that some patients had a negative length of stay (meaning the registered discharge date was before the admission date), and some patients apparently had been in treatment for up to 370 days, which seemed implausible. The average length of stay of trauma patients analysed in a publication by Chona et al. was 3.8 ± 5.4 days[4]. Since the university hospital in Lübeck treated severely injured and complicated cases, the maximum plausible length of stay was extended to the first visibly aberrant value at 130 days. Thereby, the proportion of patients with implausible extreme values was 5.98%. This value passed the threshold for an acceptable indicator. As a result, the partial weight of 1 could be included in the score calculation.

When investigating the optional data elements height and weight, the proportion of non-calculable BMI values was 53.8%, which did not pass the threshold for acceptable data quality. As a result, the partial weight of 3 could not be included in the score calculation.

The examination of the age of patients at admission showed that no patient was underage. This meant that the partial weight of 3 could be included in the score calculation.

The fracture age is displayed graphically in Figure 3, identifying extreme values using the *geom_density* function of the package ggplot2 in R Statistics[2]. It can be seen that some patients

had a negative fracture age (i.e. the registered fracture date was after the admission date), and some fractures were more than 7 days old (i.e. not complying with the procedural rules of excluding fractures >7 days old). The proportion of patients with a fracture age outside the inclusion criteria was 12%, exceeding the 5% threshold. As a result, the partial weight of 3 could not be included in the calculation. All results for calculating individual weights are summarized in Table 3[19].

The score value was calculated as follows:

Sum of individual weights ($\boldsymbol{IW}$):

$$\boldsymbol{IW = 6}$$

The sum of all specific weights ($\boldsymbol{SW}$):

$$\boldsymbol{SW = 12}$$

Values are set into the formula for scoring:

$$\boldsymbol{Score = \frac{IW}{SW} \times 100}$$

$$\boldsymbol{Score = \frac{6}{12} \times 100}$$

$$\boldsymbol{Score = 50}$$

Data quality was evaluated as moderate according to the ranking of Table 2 [19]. The fraction of cases ($\boldsymbol{n}$) taken of the 238 registered patients ($\boldsymbol{N}$), as a sample for SDV, was calculated using formula 2 and formula 3.

$$\boldsymbol{n_0 = \frac{p(1-p)}{\delta^2} \times z^2_{1-\alpha/2}}$$

$$\boldsymbol{n_0 = \frac{0,05(1-0,05)}{0,03^2} \times 1,96}$$

$$\boldsymbol{n_0 = 103}$$

$$\boldsymbol{n = \frac{n_0 \cdot N}{n_0 + N}}$$

$$\boldsymbol{n = 73}$$

Conducting the source data verification

A random sample of patients equivalent to the size of $\boldsymbol{n}$ **= 73** was drawn from the registry, using the sampling function in R Statistics. The SDV documented the percentile of discrepancy between the source and registry in the initially selected 12 items:

1. Admission date - 2.74%
2. Discharge date - 8.22%
3. Treatment date - 9.59%
4. Height and weight - 5.48%
5. Employment status - 6.85%
6. Fracture side - 9.59%
7. Number of comorbidities - 15.1%
8. Main diagnosis according to ICD-10 GM - 19.2%
9. Fracture date - 17.8%
10. Occurrence of a complication - 20.5%
11. Type of fixation - 16.4%
12. Type of reduction - 26.0%

Items 7 to 12 were further analysed using graphical methods, as they had double digit aberrations from the source.

In Figure 5, the SDV of comorbidities was further analysed, showing that, when no data was registered, patients tended to have 3 or more comorbidities (large grey circle and medium-sized pink circle). A tendency to register less comorbidities than present could be detected (second column to the right).

In Figure 6, the SDV of the coding of the main diagnoses in the registry was analysed, representing a data type with multiple elements. Besides defining an agreement between the registry and source data by setting source items as ‘same’ (multi-coloured vertical strip) when matching, patients with multiple main diagnoses were also noted (turquoise circle), suggesting the addition of a new category to the registry. For patients without an ICD-10 code in the registry (NA on the *y*-axis and grey dots horizontally at the top of the graphic), specific codes could be traced in the source data. These missing codes need to be added to the registry’s database.

It was not possible to define an agreement between the registry and source using the attribute ‘same’, for the data element ‘fracture date’. Instead, an agreement was marked with the date 1 January 1900 to make it visibly distinguishable. In Figure 7, the agreement is visible on the left part of the graphic as a thick, blue line. A fracture registered in 2016 was noted as a fresh fracture in the source data (dark blue circle bottom right). An entry error could have led to this. The same applies to a fracture in 2019 (top right circle), which was after the pilot phase. The remaining deviations seemed to lie within the range of days and to be documented at large precisely (amorphous accumulation of dots in the right area of Figure 7).

When the source data of methods of fracture reduction were investigated, further differentiation was made. When the method of reduction was registered as closed (red circles at the bottom of Figure 8), a combination or succession of methods was often found in the source. Hence, these categories were added and are suggested to be integrated into the registry. Falsely registered reduction methods were rare (no circles in the centre and left area of the graph). If no fixation information was available in the registry, the method could be assigned to the SDV afterwards (grey circles in the upper part of the graphic). The extended

differentiation, by adding additional categories, caused the double-digit percentage of deviation between the registry and source.

Testing the registration of complications

As the registration of complications was a distinguishing feature of the BFCC registry, it is graphically displayed and analysed in Figure 4. Generally, complications had been registered correctly (large green and red circle). If no information about complications was available in the registry (large grey circle), it was likely that the patient had no complications in the source data. False positive or false negative results were rare (small red and small green circle). With a confidence interval of 0.95, a sensitivity of 89.29% (71.77% to 97.73%), a specificity of 82.50% (67.22% to 92.66%) and a positive predictive value of 78.12% (64.29% to 87.63%) were calculated.

Discussion

The method of adaptive monitoring was previously published by Jacke et al. on a breast cancer query database from two 1-year episodes (1996/1997, 2003/2004)[11, 12]. In total, 877 cases were included in the study. Instead of an actual SDV, a secondary database was taken, and distributions of data were compared. This approach is suitable for large data sets, yet somehow questionable due to selection bias from the primary to the secondary database. Partly, a similar approach was used when scoring the data quality of the BFCC registry, when the length of stay was oriented on 49.778 orthopaedic and trauma patients analysed by Chona et al.[4].

Jacke et al. could reach an improvement of data quality from 51.7% to 67.7%[11], after adjusting the parameters, similar to the scoring the data quality of the BFCC fracture registry. Initially, when crude registry data was taken to calculate the score, a medium data quality with a scoring

283 result of 50 was calculated. During the SDV, it was found that the actual difference between
284 the registry and source was a mere 5.48%, in contrast to 50.8% of missing data elements.
285 Consequently, a new score was calculated, including the individual weight of 3 for the optional
286 data element of “height and weight” (Table 1). An adjusted score value of 75 was the result and
287 placed the data quality at the upper end of ‘good’ (Table 2). For future scorings of data quality,
288 a different item for the indicator ‘optional data elements’ is recommended.
289 Since no two registries use matching methods for data quality evaluation, reproducibility and
290 comparability between registries are hardly possible, which yet again shows the strength of the
291 method of adaptive monitoring. The score of the data quality has a direct consequence on the
292 sample size for SDV. On top, partly biased by the selection of parameters investigated and
293 possible modification of thresholds, attempts for comparability between data quality in
294 registries can be made. The items chosen for scoring can vary hugely from registry to registry,
295 leading to procedure bias.
296 The BFCC project chose to implement a further modification, by splitting up the indicator
297 “compliance with procedural rules” into two investigated items. The choice was made to split
298 the relatively high individual weight of 6 into two times 3 (Table 3). Hence, the compliance with
299 a legal procedural rule “patient age” at registration and a registry specific procedural rule of
300 “fracture age” could be individually taken into consideration for scoring.
301 The Anglo-American date format of ‘month/day/year’ used in the registry’s software probably
302 caused faults in both fracture age and length of stay recordings, since the date format of
303 ‘day/month/year’ is used in Germany. As the pilot phase of the registry was conducted from
304 November 2017 until February 2018, outliers in the data set are likely, as the first 12 days of
305 single digit months, like January and February, were prone to error when entering data.

The registration of comorbidities hinted towards an under-registration, as the majority of patients (52.5%) had 3 or more comorbidities. A change from a categorical variable (0, 1, 2 and >3) to a numerical variable (0, 1, 2, …, n) could improve precision.

For the systematic evaluation of registry entries, the data format in the source data evaluation needs to be chosen diligently to enable statistical analysis. Certain faults (e.g. missing ICD-10 GM codes) or necessary additional options for fixation methods (e.g. the use of an external fixator followed by internal fixation) were identified, corrections recommended, and the items added by the IT section of the BFCC project.

The graphical method *geom_count* for displaying proportions of agreement between registry data and source data (Figure 1) proved to be suitable for finding systematic faults. By using R Statistics[2] as a software tool to analyse registry data, automated reports can be created using a carefully written statistical script. This lowers the threshold for the re-evaluation of registry data and facilitates continuous improvement of the registry. As freeware, it is readily available and cost-effective for institutions to use. Furthermore, the software proved excellent when using multiple, large data sets.

Funding of the BFCC project stopped by March 2019. Implementations of suggested improvements were only carried out to a limited extent. For example, missing ICD codes were added, but the date format could not be changed by the end of the project.

Despite having a deep mathematical foundation, the method was developed with the emphasis to be used by non-mathematicians to allow for a wide application. This has been proven by this publication, as it was applied by a clinician and non-mathematician, supporting its user-friendliness and potential for broad application.

Conclusion

Scoring the data quality of a registry is a unique feature to Nonnemacher et al.'s method of adaptive monitoring and demanding in its execution, but its applicability has been proven by this publication. To tap the full potential of the method, a repeated application on an established registry would be desirable. The tested graphical method helps improving the data quality.

An outlook to monitoring data quality in the future

As the application of the method of adaptive monitoring has yet only been published for two registries, possibilities for further research are vast. Its application on different projects could further test its reliability, with the aim to make it the gold standard for evaluating the data quality of registries.

To limit transfer and human error and safe time, an automated data capture should be considered in the future. The excessive manpower needed to acquire sufficient amounts of data for the BFCC registry was outdated. Solving this issue was subject of a different branch of the BFCC project, focusing on import and export solutions from the Hospital Information System (HIS) to the registry's database. It could not be fully executed by the end of the project for reasons of software and data format incompatibilities. It is advisable for any new registry to meticulously care for data formats prior to setting it up or evaluating the data quality.

A shortcoming of the project was a selection bias, as only patients able to consent were included in the registry. As orthopaedic and trauma departments often deal with fragility fractures of older and not contractually capable patients, a bypass through an opt-out system, as used in Scandinavia[10, 17] or the Netherlands[21], would facilitate including patients. A drop-out rate caused by this was unfortunately not tracked and is suggested to be recorded by future projects.

354 List of Abbreviations

355 BFCC – Baltic Fracture Competence Centre

356 BMI – Body Mass Index

357 EU – European Union

358 ICD-10 GM – International Statistical Classification of Diseases and Related Health Problems,

359 10th edition, German Modification

360 IT – Information Technology

361 IW – Individual Weights

362 SDV – Source Data Verification

363 SW – Sum of Weights

364 UKSH – Universitätsklinikum Schleswig- Holstein

365

366

Declarations

Ethics approval and consent to participate

The research project was approved by a suitably constituted Ethics Committee of the University of Lübeck on 9 October 2017, reference number 17-267. All study participants gave informed consent. Consent to participate was obtained in written form.

Consent for publication

Not applicable

Availability of data and material

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

Not applicable

Authors' contributions

All authors read and approved the final manuscript.

JF – mainly contributed to writing the manuscript and statistically analysing the data as well as interpreting results and drawing conclusions

APS – contributed as leader of the BFCC Project by helping to find statistical methods to analyse data quality of the registry and proofreading the manuscript

AG – contributed by enrolling patients into the study

AW – contributed to the availability of data from the hospital information system and statistical analysis

GH – contributed by enrolling patients into the study

Acknowledgements

Not applicable

References

1. Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V, TMF, 2018, http://www.tmf-ev.de/Home.aspx, German
2. R Statistics, 2018, Statistical Software, Version 3.5.1 + RStudio, Version 1.1.45, https://www.r-project.org/foundation/
3. 2019, PubMed, 8600 Rockville Pike, Bethesda MD, 20894 USA, https://www.ncbi.nlm.nih.gov/pubmed/
4. Chona, D., N. Lakomkin, C. Bulka, I. Mousavi, P. Kothari, A.C. Dodd, M.S. Shen, W.T. Obremskey, and M.K.J.I.o. Sethi, *Predicting the post-operative length of stay for the orthopaedic trauma patient.* International Orthopaedics, 2017. **41**(5): p. 859-868.
5. Flachenecker, P., U. Zettl, U. Götze, J. Haas, S. Schimrigk, W. Elias, M. Pette, M. Eulitz, M. Hennig, and J.J.D.N. Bertram, *MS-Register in Deutschland—Design und erste Ergebnisse der Pilotphase.* Der Nervenarzt, 2005. **76**(8): p. 967-975.
6. Gjerstorff, M.L.J.S.j.o.p.h., *The Danish cancer registry.* Scandinavian Journal of Public Health, 2011. **39**(7_suppl): p. 42-45.
7. Google, L. *Google Scholar*. 2019 September 9 2018]; Available from: scholar.google.com.
8. Graves, S.E., D. Davidson, L. Ingerson, P. Ryan, E.C. Griffith, B.F. McDermott, H.J. McElroy, and N.L.J.M.J.o.A. Pratt, *The Australian orthopaedic association national joint replacement registry.* 2004. **180**(5): p. S31.
9. Helgesson, Ö., C. Bengtsson, L. Lapidus, C. Merck, and P.J.S.j.o.s.m. Sparén, *Malignant Disease Observed in a Cohort of Women A Validation of Swedish Cancer Registry Data.* Scandinavian Journal of Public Health, 1994. **22**(1): p. 46-49.
10. Hillert, J. and L. Stawiarz, *The Swedish MS registry–clinical support tool and scientific resource.* Acta Neurologica Scandinavica, 2015. **132**: p. 11-19.
11. Jacke, C., M. Kalder, M. Koller, U. Wagner, and U. Albert, *Systematische Bewertung und Steigerung der Qualität medizinischer Daten.* Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz, 2012. **55**(11-12): p. 1495-1503.
12. Jacke, C.O., M. Kalder, U. Wagner, and U.-S. Albert, *Valid comparisons and decisions based on clinical registers and population based cohort studies: assessing the accuracy, completeness and epidemiological relevance of a breast cancer query database.* BMC research notes, 2012. **5**(1): p. 700.
13. Juto, H., M. Möller, D. Wennergren, K. Edin, I. Apelqvist, and P. Morberg, *Substantial accuracy of fracture classification in the Swedish Fracture Register: Evaluation of AO/OTA-classification in 152 ankle fractures.* Injury, 2016. **47**(11): p. 2579-2583.
14. KAIROS, 2018, Centraxx Biobank, https://www.kairos.de/produkte/centraxx/
15. Larsen, I.K., M. Smastuen, T.B. Johannesen, F. Langmark, D.M. Parkin, F. Bray, and B. Moller, *Data quality at the Cancer Registry of Norway: an overview of comparability, completeness, validity and timeliness.* Eur J Cancer, 2009. **45**(7): p. 1218-31.

16. Laustsen, J., L. Jensen, and A. Hansen, *Accuracy of clinical data in a population based vascular registry.* European Journal of Vascular and Endovascular Surgery, 2004. **27**(2): p. 216-219.
17. Lehtinen, M., H.M. Surcel, K. Natunen, E. Pukkala, and J. Dillner, *Cancer Registry follow-up for 17 million person-years of a nationwide maternity cohort.* Cancer medicine, 2017. **6**(12): p. 3060-3064.
18. Meling, T., K. Harboe, A.J. Arthursson, and K. Soreide, *Steppingstones to the implementation of an inhospital fracture and dislocation registry using the AO/OTA classification: compliance, completeness and commitment.* Scand J Trauma Resusc Emerg Med, 2010. **18**: p. 54.
19. Nonnemacher, M., D. Nasseh, and J. Stausberg, *Datenqualität in der medizinischen Forschung*. Vol. 4. 2014: Medizinisch Wissenschaftliche Verlagsgesellschaft.
20. Pipino, L.L., Y.W. Lee, and R.Y. Wang, *Data quality assessment.* Communications of the ACM, 2002. **45**(4): p. 211-218.
21. Rakhorst, H.A., M.A. Mureau, R.D. Cooter, J. McNeil, M. van Hooff, R. van der Hulst, J. Hommes, M. Hoornweg, L. Moojen-Zaal, and P. Liem, *The new opt-out Dutch national breast implant registry–Lessons learnt from the road to implementation.* Journal of Plastic, Reconstructive & Aesthetic Surgery, 2017. **70**(10): p. 1354-1360.
22. Schneemann, I. *BFCC- Project Webpage*. 2018; Available from: http://bfcc-project.eu/project.html.
23. Veen, E.J., M.L. Janssen-Heijnen, E. Bosma, M.A. de Jongh, and J.A. Roukema, *The accuracy of complications documented in a prospective complication registry.* Journal of Surgical Research, 2012. **173**(1): p. 54-59.
24. WebMD. *Medscape*. 2019 September 9th 2018]; Available from: medscape.com.
25. Zaltman, G. and C. Moorman, *The importance of personal trust in the use of research.* Journal of Advertising Research, 1988. **28**(5): p. 16-24.

Table Titles

Table 1. Items, levels, and indicators for scoring of data quality. Values are adjusted to the BFCC project, yet orient on recommended values by Nonnemacher et al.

Table 2. Delta value in relation to score result and data quality.

Table 3. Results for the calculation of individual weights. Values are adjusted to the BFCC project, yet orient on recommended values by Nonnemacher et al.

Figure Legends

Figure 1. Arbitrary example to introduce the statistical method for source data verification. The perception of the hospital personnel (y-axis) versus their actual function (x-axis). Prop = proportions

Figure 2. Length of stay of all patients in the registry.

Figure 3. Fracture age (in days) of all patients in the registry.

Figure 4. Occurrence of complications in all patients. Prop = proportions.

Figure 5. Source data verification on the registered number of comorbidities. Prop = proportions.

Figure 6. Graphical analysis of International Classifications of Diseases (ICD) coding in source data verification. Prop = proportions.

Figure 7. Source data verification of fracture date. Prop = proportions. Instead of “same’’, the date 1 January 1900 was selected.

Figure 8. Source data verification (SDV) of reduction method. Prop = proportions.

Formulas

Formula 1: Calculating the score result

Formula 2: Calculating the unadjusted case number

Formula 3: Calculating the adjusted case number

Figures

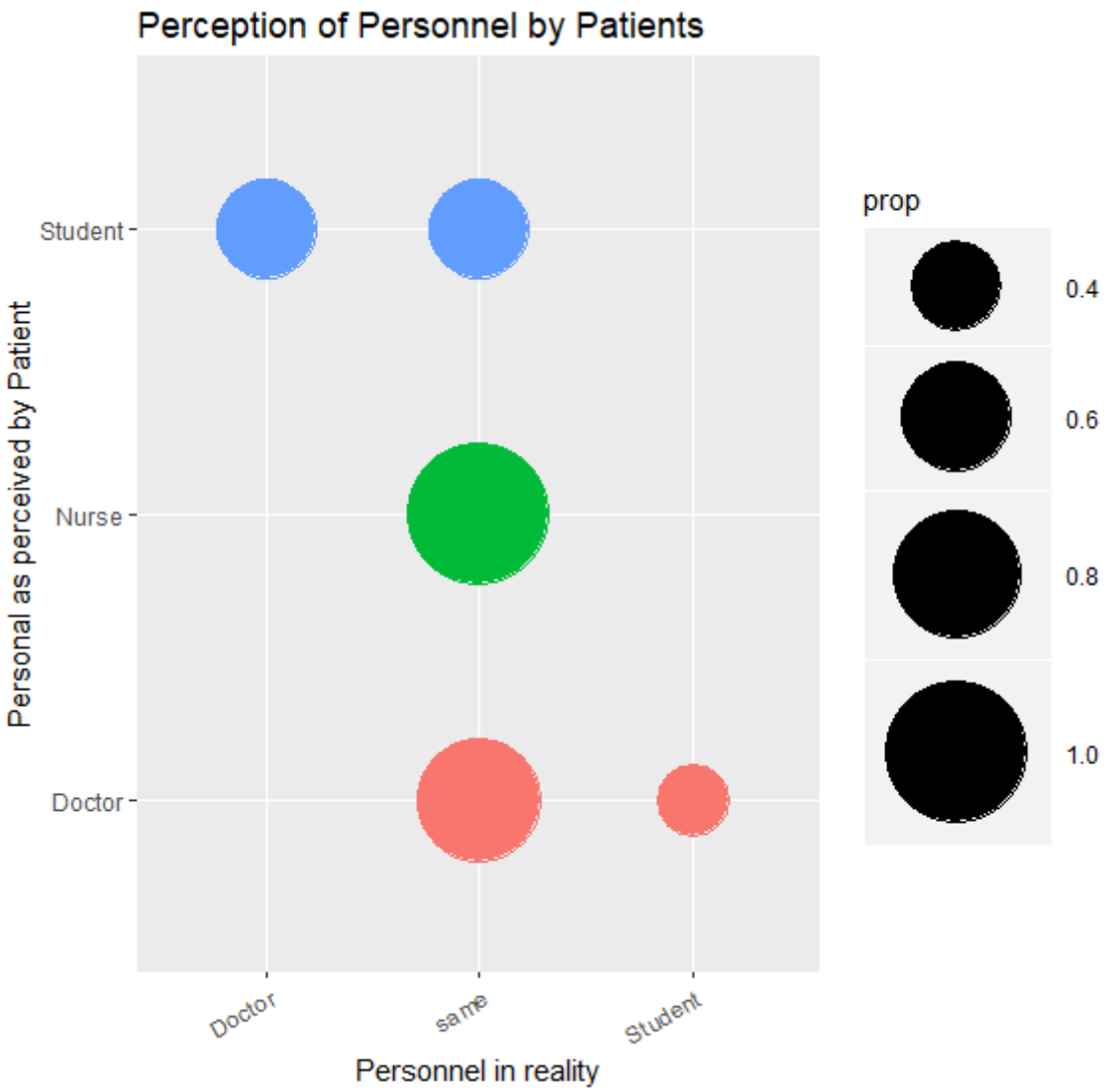

Figure 1

Arbitrary example to introduce the statistical method for source data verification: The perception of the hospital personnel (y-axis) versus their actual function (x-axis). Prop = proportions

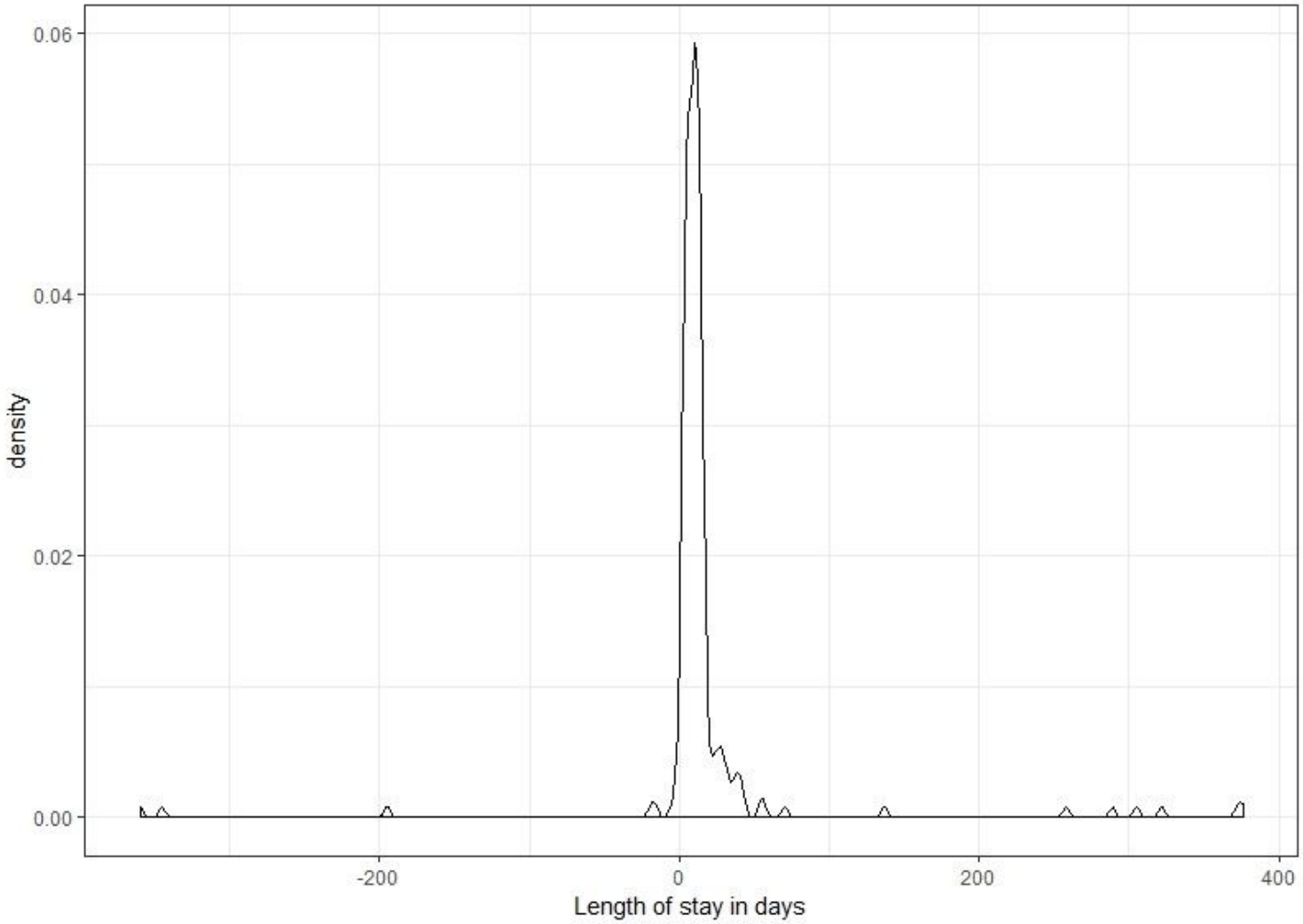

Figure 2

Length of stay of all patients in the registry.

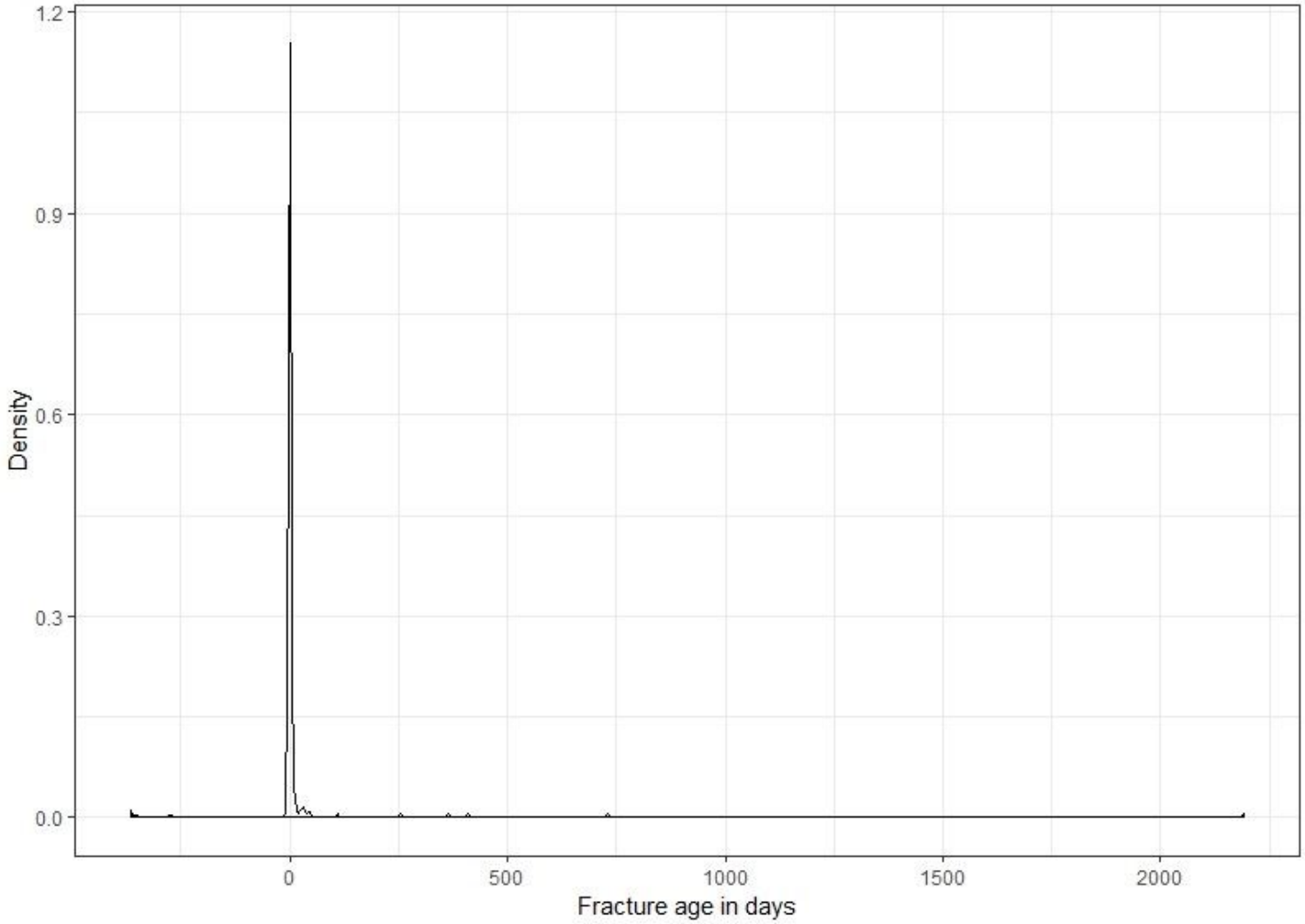

Figure 3

Fracture age (in days) of all patients in the registry.

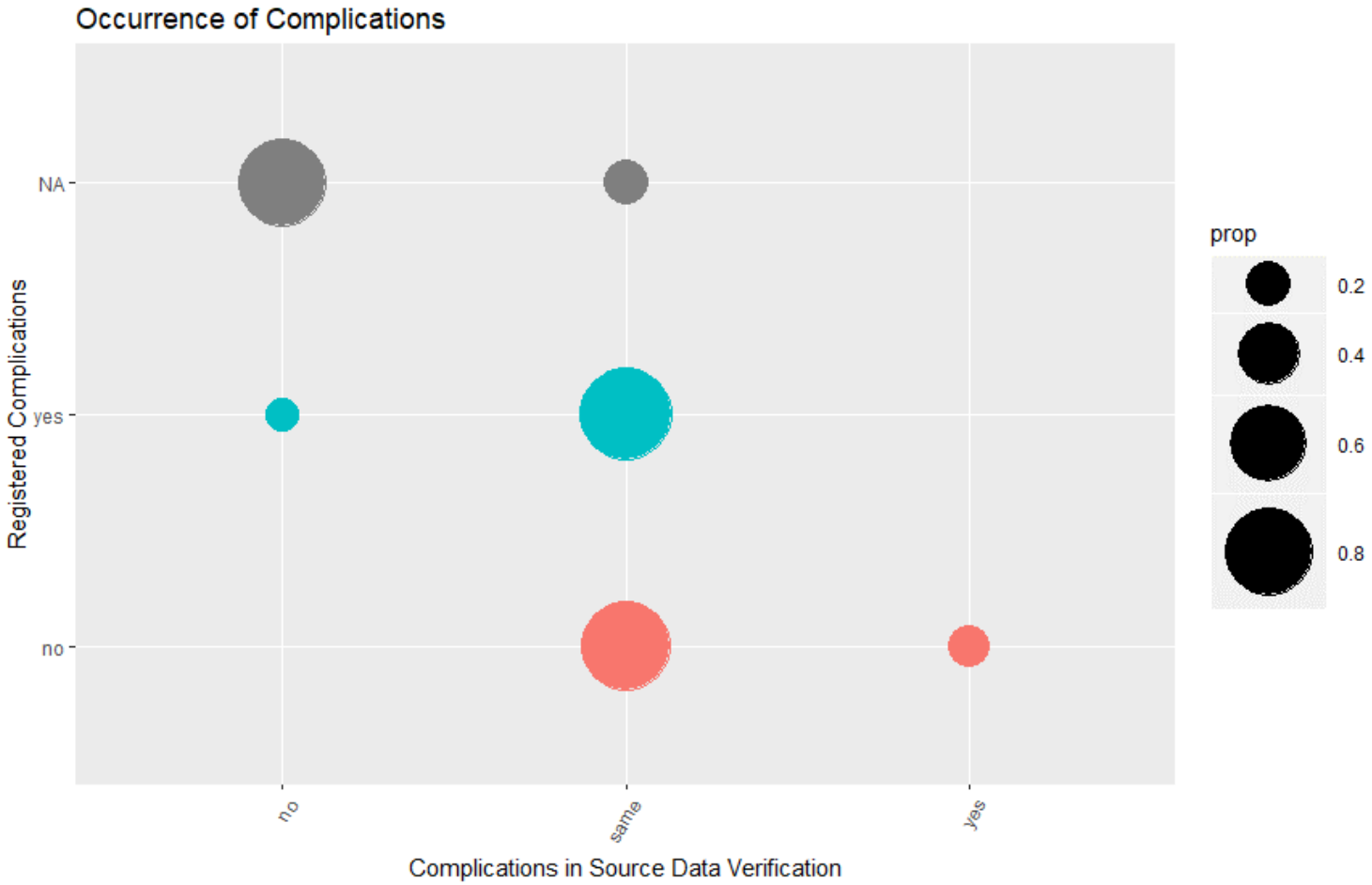

Figure 4

Occurrence of complications. Prop = proportions.

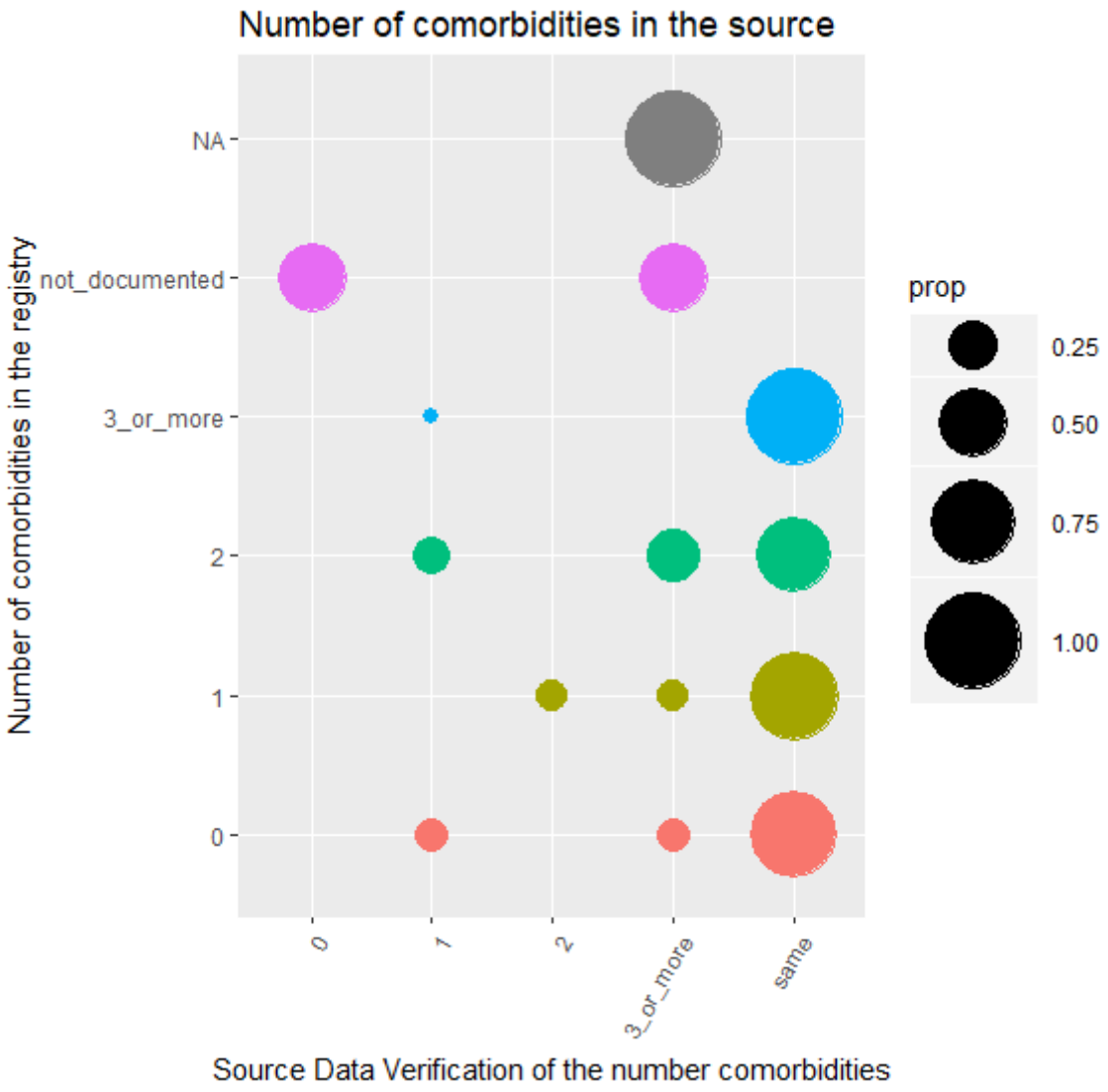

Figure 5

Source Data Verification on the registered number of comorbidities. prop = proportions

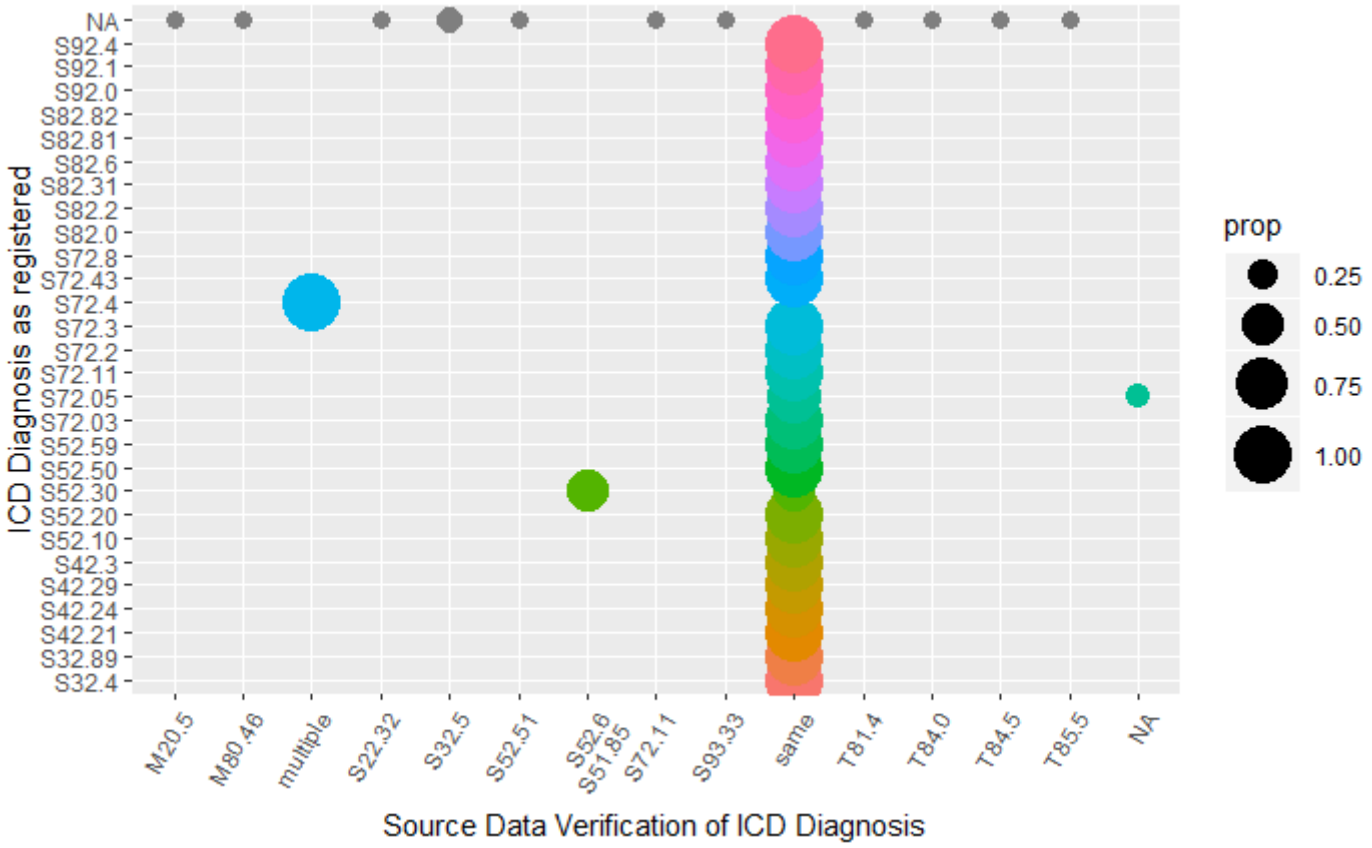

Figure 6

Graphical analysis of ICD (International Classifications of Diseases) coding in source data verification. prop = proportions

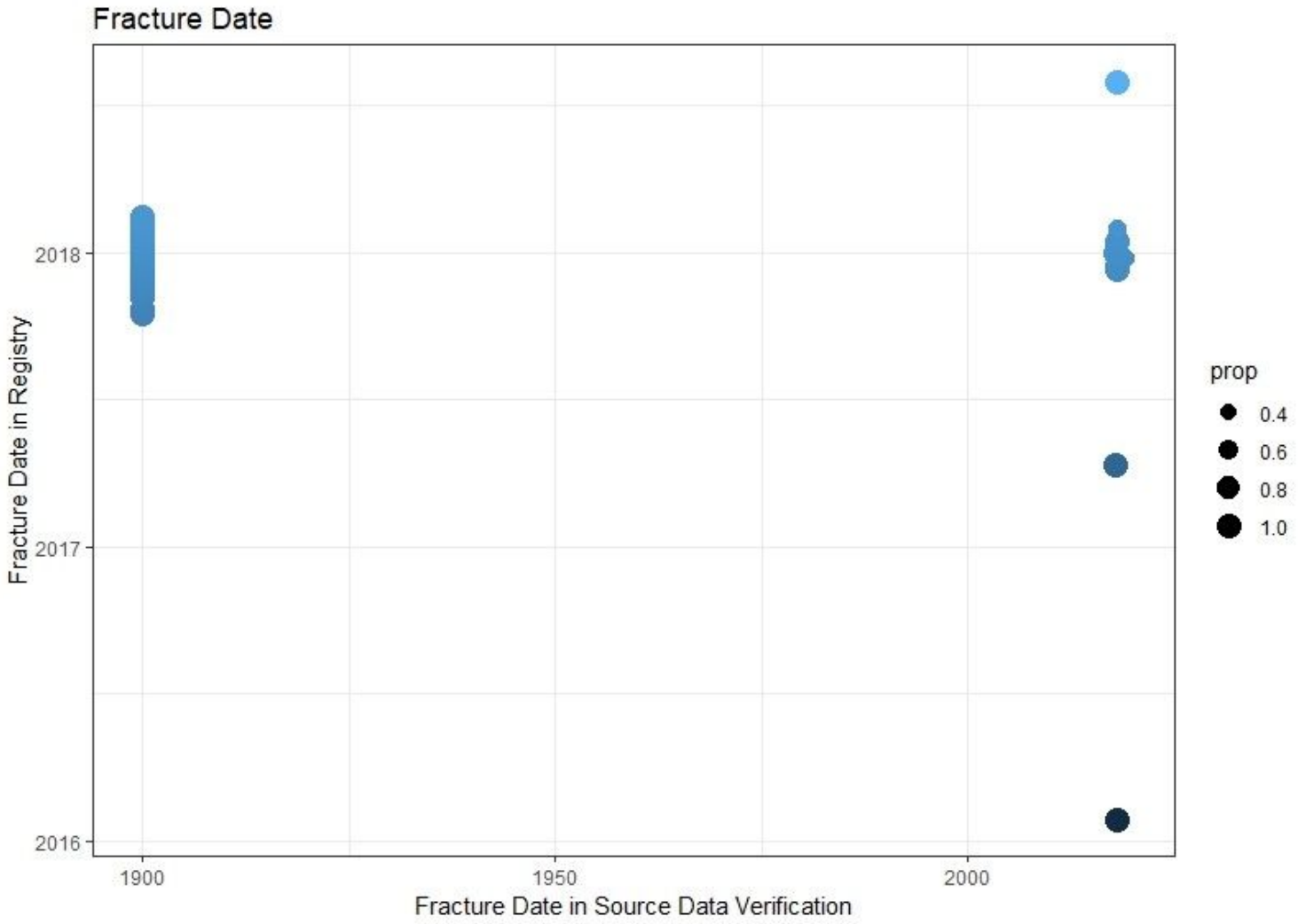

Figure 7

Source data verification of fracture date. Prop = proportions. Instead of ‘same”, the date 1 January 1900 was selected.

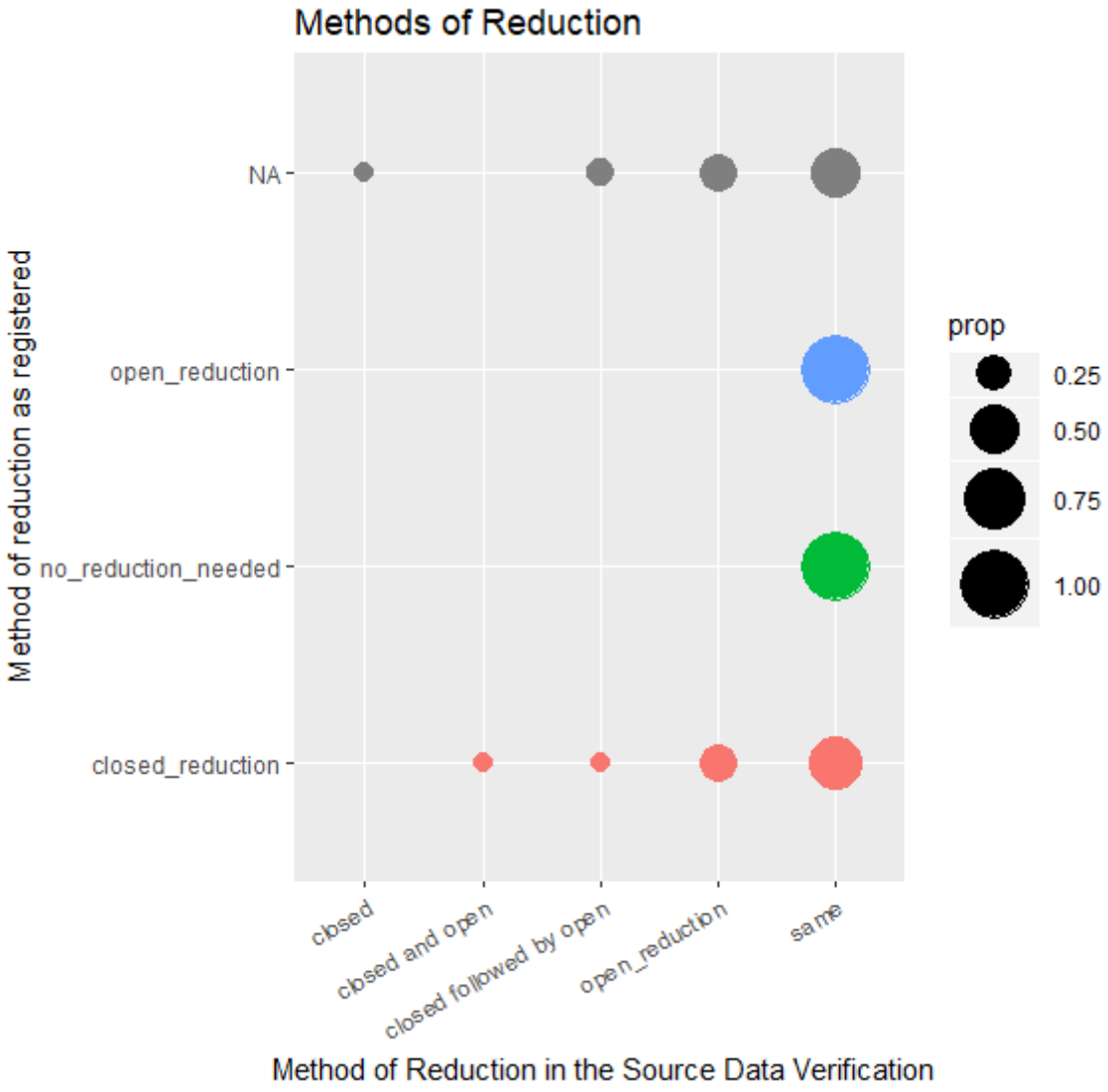

Figure 8

Source data verification (SDV) of reduction method. Prop = proportion

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table1.JPG
- Table3.JPG
- Table2.JPG