ORIGINAL ARTICLE

Completeness and accuracy of data in spine registries: an independent audit-based study

Bernhard Meyer[1] · Ehab Shiban[1] · Lucia E. Albers[1] · Sandro M. Krieg[1]

Abstract

Purpose Clinical registries are used for quality management and clinical research. Due to the importance and implications of both aims, completeness and high quality of data are of paramount importance. However, this remains uncertain, as none of these registries have implemented independent monitoring. The aim of this study was to determine the accuracy and completeness of registry data o the example of the German Spine Society (DWG) registry.

Methods In a prospective study, audits by a board-certified neurosurgeon were conducted at certified spine centers with mandatory registry input, a setting comparable to most existing registries worldwide. A 2-week period was analyzed, and any discrepancy between patients' charts and the registry entry was evaluated. A median of 31 items per patient was evaluated including completeness and accuracy of data.

Results Out of 17 centers willing to participate, 4 were still lacking any data entries. Even in the remaining 13 centers eligible for audits, 28.50% (95%-CI = [22.46–34.55]) of entries were finalized only after the audits were announced. Only 82.55% (95%-CI = [79.12–85.98]) of surgeries were documented, and on average 14.95% (95%-CI = [10.93–19.00]) of entries were not accurate with a wide variation (range; 6.21–27.44%) between centers. Aspects for improvement of the situation were identified.

Conclusion Due to the high inaccuracy, the high number of centers lacking mandatory entries at all and the number of false entries, these data alert us to advocate unannounced audits and further measures to improve the situation. Data should not be used for the time being, since wrong conclusion will be drawn.

Graphic abstract

These slides can be retrieved under Electronic Supplementary Material.
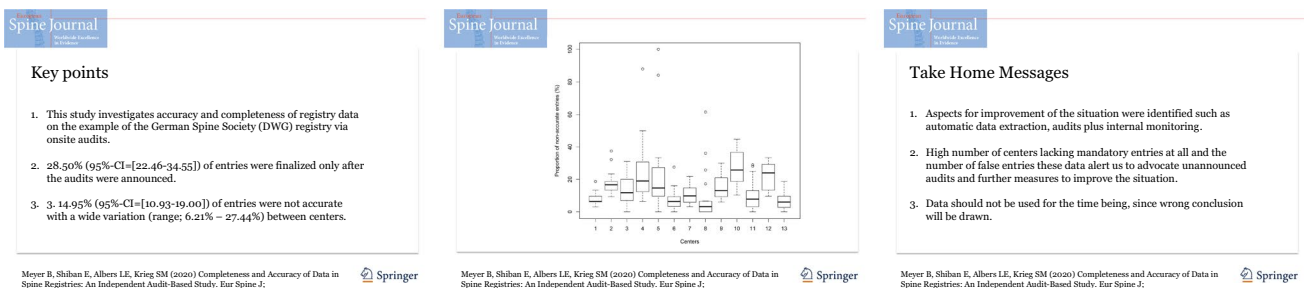
Keywords Spine · Registry · Scientific value · Evidence-based medicine · Quality management

✉ Sandro M. Krieg
Sandro.Krieg@tum.de

Extended author information available on the last page of the article

List of symbols

DWG	German Spine Society
RCT	Randomized controlled trial
CI	Confidence interval

Introduction

Registries were designed to acquire comprehensive data on treatment effects and natural course of diseases [1]. Particularly in spine surgery, registries have been adopted very early. The first structured and large-scale spine registry was the Swedish Spine Registry; Spine Tango was the first international [2, 3]. Thus, we look back on more than two decades of experience with data originating from registries, which are increasingly used for clinical science and quality management.

Randomized controlled trials (RCTs) enable us to gain level I evidence for treatment strategies but have the disadvantage of well-defined subgroups as their target of research and a standardized environment that does not reflect clinical reality.

Registries have the hypothetical advantage to acquire data from a complete population with a specific disease treated in a real-world setting. Thus, clinical studies with data from registries are considered to ideally complementing RCTs or vice versa.

Data originating from clinical registries become also increasingly important today for quality management purposes to all stakeholders of our healthcare systems, e.g., health insurances, healthcare providers, politicians, professionals and their respective scientific societies. For example, new legislation in Europe requires that medical products such as spinal implants can be individually followed and are correlated with clinical data. Only registries can fulfill these requirements. It is further to be expected or already in place that decisions regarding reimbursement of treatment or even sanctions against hospitals and/ or surgeons are derived from comparative registry data.

Because these two purposes are highly important and may have far reaching consequences, data quality of registries itself with regard to completeness and accuracy becomes of utmost importance.

In strong contrast to this stands the fact that few spine registries have external monitoring, which would be a first prerequisite to assure data quality.

Another obstacle for almost all registries with exception of the Scandinavian ones is that most national data protection laws require strictly anonymized data entries, which prohibits any relevant and accurate follow-up for very important markers such as reoperation rates or 30-day readmissions.

Thus, the aim of this study was to investigate the accuracy and completeness of data from registries as they are today via the example of the German Spine Society Registry. Since we are convinced that the findings can be extrapolated to most other registries, we consider our message important beyond the spine community.

Methods

Ethics approval and consent to participate

The local ethics committee of our university confirmed the study (registration number 42/16). All patients gave written informed consent to the registry.

Literature search prior to the study

We performed a search for currently ongoing registries and their data quality via Medline, Google, clinicaltrials.gov, and inquiries to societies. We included registries with focus on scientific, socioeconomic and healthcare provider perspectives, which are used and acknowledged by national scientific societies or governmental organizations. For the purpose of this article, we did not consider registries of scientific organization and/or universities/ hospitals, focusing on one specific spinal disease or one scientific question. Search criteria were "audit," "registry," "accuracy," "data completeness," and "monitoring."

Study design

STROBE guidelines were used to execute this study in accordance with cohort studies.

The German Spine Society (Deutsche Wirbelsäulengesellschaft, DWG) was contacted and confirmed the institutional certification of 23 German centers. Institutional certification is linked to mandatory data input into the DWG registry.

All the centers were contacted, and the participating centers were chosen according to Fig. 1. The minimum mandatory requirement upon certification was to document in the registry all inpatient data of patients undergoing spine surgeries. For the audit visit, a 2-week period was defined which was 2 months before the initial contact to the centers, in which all patients, which had had spine surgery, were analyzed. In a prospective study, design audits by a board-certified neurosurgeon were then planned and conducted at all participating centers. The percentage of surgeries that were not documented in the registry at the time of the first contact concerning the study as well as the overall rate of undocumented cases were analyzed. Moreover, any discrepancy between the patients' charts and the registry entry was detected and analyzed at the audit visit. A mean of 30.95 ± 3.05 items per patient (median 31, range 0–40) and a mean of 20.92 ± 7.96 (median 20, range 11–38) cases per center were analyzed.
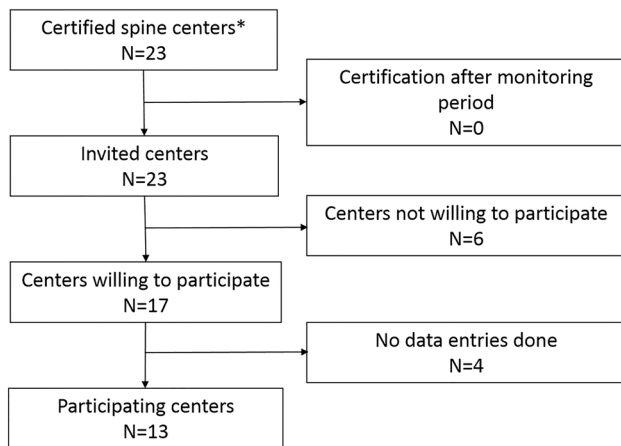
* Registry entry for all operated spine cases is mandatory

Fig. 1 Flowchart: This chart shows the exclusion of several centers due to the lack of data entries in a considerable number

Variables

The analyzed items per case included main diagnosis, type of fracture, type of degeneration, classification of spondylolisthesis and osteoporosis, affected levels, type of tumor, previous surgeries, type and duration of symptoms, date of surgery, surgeon, implants, type of surgery, postoperative course, complications, etc. Table 1 provides a comprehensive overview.

Statistical analysis

The SPSS v.24 (IBM Inc., Armonk, New York) statistical software package and R statistical software (version 3.1.0; https://www.r-project.org/), Vienna, Austria were used for the analysis. The objective of our study was to deliver data on the general accuracy of data originating from a large national registry per se and on the other hand to identify potential factors leading to an increased rate of wrong entries depending on the item.

Results are presented as proportions or means with 95% confidence intervals (CI). Wilcoxon–Mann–Whitney statistics and Kruskal–Wallis statistics were used to test for differences in accuracy rates of the entries depending on a) monitoring of the data entry and b) the person performing the data entry were entered.

Results

Overall data completeness

Out of 23 certified spine centers contacted, 17 centers were willing to participate, but 4 were still lacking *any* data

entries. Even in the remaining 13 centers eligible for audits, 28.50% (95%-CI = [22.46–34.55]) of entries were finalized only after the audits were announced. Thus, 71.50% (95%-CI = [65.45–77.54]; range between centers: 18.18–100%) of all operated patients were completely entered into the registry at the time of the first contact (Fig. 2).

At the time of the visit only, 82.55% (95%-CI = [79.12–85.98]; range between centers: 20.41–100%) of surgeries were documented in the registry although all centers are required to enter all operated spine cases without exception (Fig. 3).

Incorrect entries per center

On average, 14.95% (95%-CI = [10.93–19.00]) of entries were not accurate with a variation (range: 6.21–27.44%) between centers (Fig. 4). The setup also varied largely in the centers. While the registry entries were done by the operating surgeon in 10 centers, they were done by a study nurse in three centers. Looking at the influence of the person entering data, there is a significant difference with fewer errors if the surgeon (12.02% [10.47–13.58]) performs the data entry compared to a study nurse (15.65% [14.22–17.68]) (Fig. 5; $p = .0004$).

Table 1 shows the percentage of wrong registry entries per item. Wrong entries per item varied largely and were between .4 and 50% of the entries. It is worth mentioning that complications as one of the most useful data from large registries showed a failure rate of 50%.

Influence of internal monitoring

In four centers, no internal monitoring was performed. Two centers perform an internal check whether the patient was entered into the registry (quantity) but without any content control while there is a quality monitoring in seven centers, which is either done by an attending or by a study nurse. Yet, the percentage of wrong entries did not differ significantly for all three groups (no monitoring: 10.10% [8.75–11.46]; quantitative monitoring: 17.83% [11.96–23.67]; qualitative monitoring: 14.72% [12.29–17.15], $p = .008$; Fig. 6a). Comparing "no qualitative monitoring" (12.00% [10.16–13.84]) vs. "qualitative monitoring" (14.72% [12.29–17.15]), we also do not observe any significant difference ($p = .084$; Fig. 6b).

Discussion

Added value of this study

Our study will raise awareness on the general shortcomings of registries even if supervised by large societies and

Table 1 Analyzed items

Item	% of wrong entries	No. of wrong entries	No. of entries
Diagnosis			
Main diagnosis	12.4	33	266
Primary type of degeneration	17.7	27	153
Type of degeneration	23.9	28	117
Type of deformity	37.5	6	16
Main etiology	16.7	1	6
Type of fracture/trauma (AO A, B, C)	10.0	4	40
Type of fracture/trauma (AO 1, 2, 3)	11.8	2	17
Pathological fracture due to	12.5	3	24
Age of fracture	0	0	32
OF classification	25.0	3	12
Type of spondylolisthesis	0	0	16
Degree of spondylolisthesis	23.1	6	26
Infection details	25.0	1	4
Affected structure	0	0	4
Type of tumor	10.5	2	19
Location of tumor	15.8	3	19
Reason for reoperation	31.4	11	35
Affected levels	9.9	26	263
No. of levels	2.6	7	266
further diagnoses	19.8	45	227
No. of previous surgeries at same/adjacent level	6.4	17	264
No. of previous surgeries at other level	3.4	9	263
Duration of symptoms	11.9	31	261
Operation			
Date of surgery	0.4	1	268
Surgeon	21.6	56	259
Assistant	37.5	51	136
Therapeutic goals	32.2	86	267
Implants	34.0	91	268
Anterior approach	1.1	3	268
Posterior approach	12.7	34	267
ASA classification	24.7	65	263
Surgical details			
Decompression	31.3	83	265
Levels of decompression	10.9	23	211
Fusion type	8.3	22	265
Levels of fusion	9.9	8	81
Fusion material	10.4	16	153
Rigid instrumentation	4.9	13	268
Levels of instrumentation	10.0	10	100
Deformity correction	4.1	11	266
Levels of deformity correction	8.0	2	25
Movement preserving surgery	4.1	11	267
Levels of movement preserving surgery	0	0	6
Percutaneous measures	0	0	1
Others	17.2	45	262
Levels of percutaneous measures	10.3	4	39
Intraoperatively unintended events	4.1	11	266
Reintervention during same surgery	35.0	7	20

Table 1 (continued)

Item	% of wrong entries	No. of wrong entries	No. of entries
Intraoperatively general complications	2.1	5	242
Intraoperatively surgical complications	50.0	1	2
Inpatient stay			
Postoperative surgical complications	10.6	28	264
General complications	4.9	13	264
Reintervention	29.3	17	58
Inpatient stay	8.0	9	112
Status of surgical complications	0	0	1
Status of therapeutic goals	0	0	1
Certification specificities			
Entity	9.6	25	260
Points	8.1	21	260

This table provides all analyzed items which were examined for each patient depending on the type of disease treated. Moreover, the number of analyzed patient with each item and the percentage of wrong entries for each item are provided

Fig. 2 Completed entries upon first contact: This chart shows the percentage of completely entered cases into the registry in relation to all entered patients per center at the time of the first contact for the study
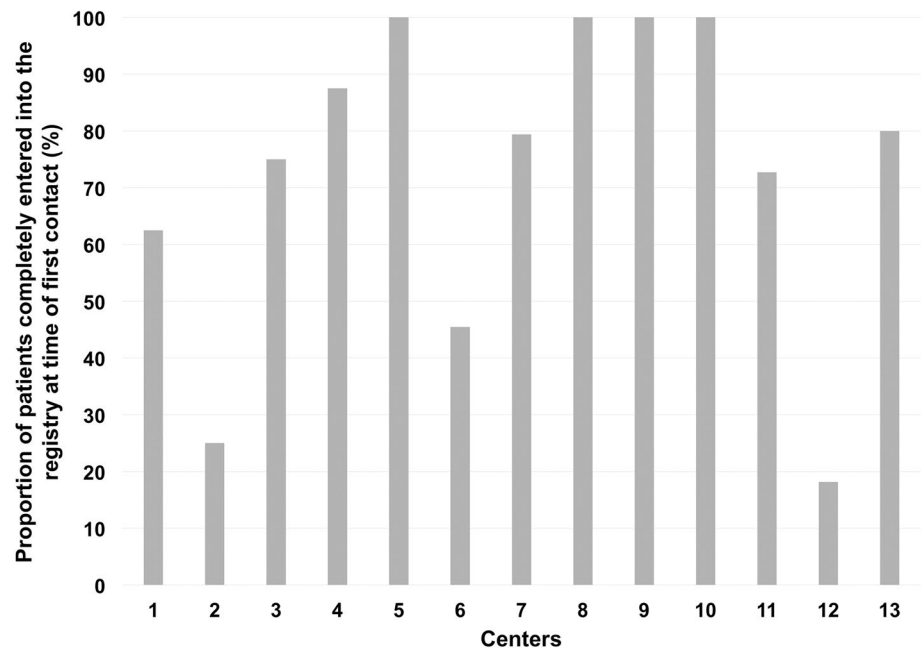
will help to improve future registry design and healthcare decisions. It is particularly valuable because it is actually the first study to determine the accuracy and completeness of a large nationwide registry.

In the light of today's socioeconomic and health political influence of data gained from large national registries, this study provides profound evidence that the design and structure of most current registries does not provide reliable data sufficient for political, medical, or economic decision making.

General aspects

Our results indicate that the data of the DWG registry are neither suitable for clinical science purposes nor quality management for the time being. Compliance with mandatory data input was merely moderate. Completeness of data was not satisfactory on average, but especially inacceptable in a significant fraction of individual centers despite the fact that only the minimum data set was required, i.e., a reduced inpatient data set without follow-up. Accuracy of data was

Fig. 3 Missing of whole cases: This chart shows the percentage of data entries in relation to all operated patients per center
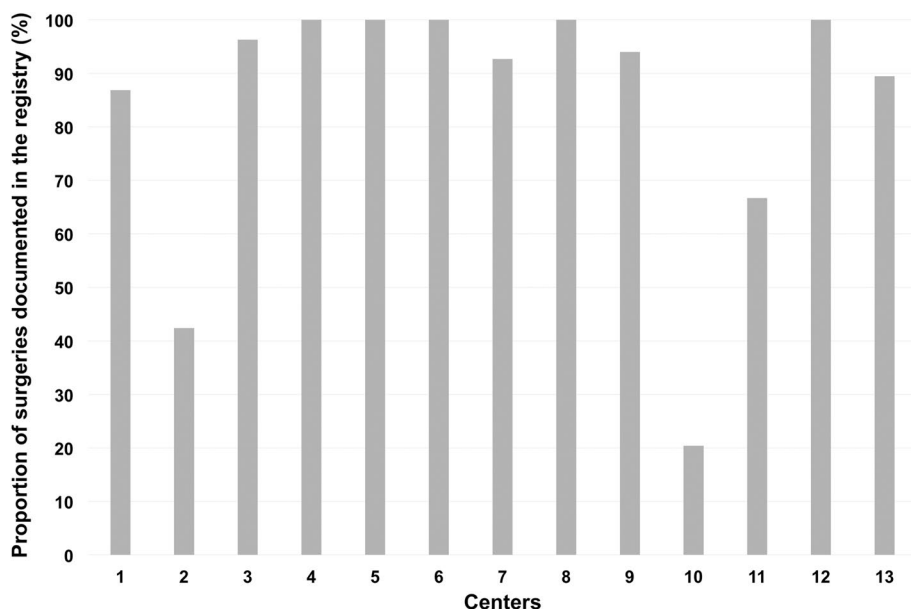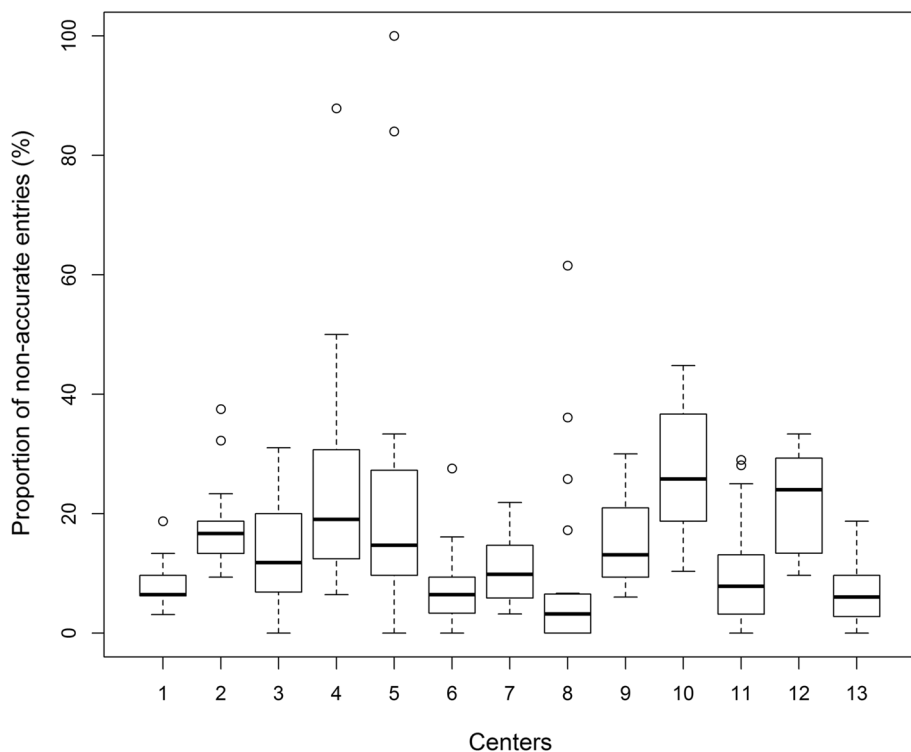

Fig. 4 Wrong entries per center: This box plot shows the percentage of wrong data entries of the registry versus the patient charts in relation to all data entries per center

also not satisfactory, again with a wide variation between centers. Internal non-mandatory monitoring did neither influence completeness nor accuracy of data.

Data accuracy and audits

There are various aspects, which could improve data quality and should also be pursuit such as a strict definition of validated key metrics for each disease, enabling automated

data extraction, and providing adequate financing for sufficient resources.

The DWG registry examined in this study is based on the Spine Tango platform. As mentioned above, the reliability and usefulness of the results are inherently dependent on the quality and completeness of data being entered into the registry. RCTs undergo regular monitoring visits to ensure accurate and complete data entry, but most of the current registries fail to have mandatory comprehensive-auditing
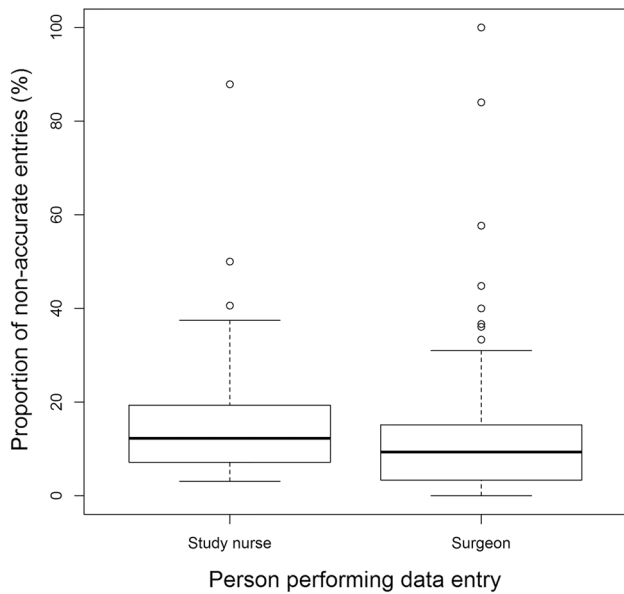
Fig. 5 Entering person: This box plot shows the influence of the entering person: there is a significant difference with fewer errors if the surgeon performs the data entry ($p = .0004$)

procedures. The problems identified in this registry seem therefore inherent to the majority of spine registries and probably beyond.

Of course, in order to prevent selection bias and also to account for the actual variances between centers concerning outcomes, multivariate analyses are required adjusting for covariates [4]. But multivariate analyses can only handle known covariates or bias. If these are unknown or not recorded by the registry, we have no chance to clear the registry data from these influences potentially leading to wrong or unrepresentative results. Thus, the finer granularity is enabled by the required items per patients, the more accurate and the more unlikely that the data set suffers from unknown bias. However, fine granularity is more elaborate for the centers, thus leading to incomplete data sets. Regarding Table 1, there are items that are more prone to incorrect entries than others, and it is important to control these issues and change the items themselves accordingly.

It is important to say that the large variance between centers (Fig. 4) also points out that we need to differentiate between inherent problems of a registry itself causing low data accuracy and completeness and culture in the participating centers themselves.

Likewise, a large amount of missing data will also limit the value of the registry data. Thus, direct and real-time internal monitoring of data quality and integrity is crucial when designing and planning a registry. Moreover, without an active external monitoring or auditing process, the danger of selective reporting/entry is obvious. Particularly, the reported rate of complications seems extremely low in our study (Table 1). Bias can occur by factors such as omitting whole groups of high-risk patients or liberal interpretation of what constitutes as a complication. Both factors are known to occur if no rigorous guidelines for reporting are established and ultimately controlled. To our knowledge, this is not the case for any European spine registry at least.
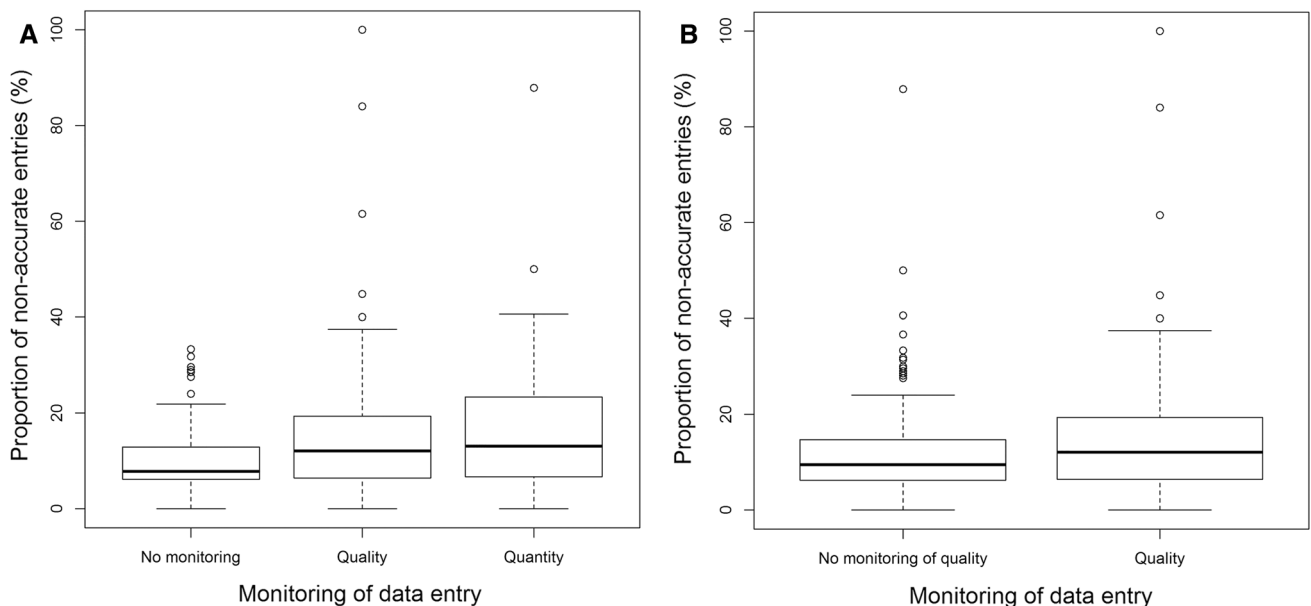

Fig. 6 Internal monitoring: This box plot demonstrates the percentage of wrong entries depending on the presence or absence of internal monitoring. **a** For all three groups (no, quantity, and quality; $p = .008$); **b** "no qualitative monitoring" versus "qualitative monitoring" ($p = .084$)

Follow-up versus data protection

Another problem not examined in this study is the lack of relevant follow-up. Two aspects exist. First, follow-up rates tend to be low in general, but for clinical research, collecting follow-up data remains crucial. It is known that any loss of follow-up larger than 20% of the enrolled cases can lead to a significant bias in the statistical evaluation [5]. In a study by McGirt et al., the follow-up rate in spine registries was found to be between 22 and 79% [6]. Thus, even the registry with the best follow-up rate was worse than minimally recommended one. In our study, only 71.50% (95%-CI = [65.45–77.54]; range between centers: 18.18–100%) of all operated patients were completely entered into the registry at the time of the first contact (Fig. 2); and this even only reflects follow-up until discharge.

This issue relates to the current state of legislation regarding data protection in most countries and is problematic for quality management. To calculate some of the key indicators for quality of care such as 30-day readmission or reoperation rates, a patient would need to be traced, irrespective of the hospital taking care of the complication. Particularly in urban areas with a high density of specialists, a significant number of patients will seek a second opinion for problems occurring after the index surgery. Identification of a given patient for longitudinal follow-up is crucial. Outside Scandinavia, however, this is not possible in any other country we are aware of, which means that registries in those countries provide data that will clearly underestimate important quality indicators.

Further approaches to improve data quality

In order to find further solutions for the imminent problems raised by this study, awareness among peers teaching on the potentially crucial implications of registry data but also possible financial incentives could be used. Changing culture is another part.

Yet, when looking at the involved centers, those were the first certified spine centers in Germany and not only by this fact but also by personal knowledge highly motivated professionals, in part leaders in the field and well motivated. Thus, blaming lack of motivation, culture of education on the topic might not serve as complete solution. Other issues, such as shortness of staff, amount of required data and lacking usability of the registry software itself, should also be taken into account.

Conclusion

Due to the high inaccuracy, the high number of centers lacking mandatory entries at all and the number of false entries these data alert us to advocate unannounced audits and other measures to improve the situation such as financial incentives and education on the benefits and consequences of such data. The current data should not be used for the time being, since wrong conclusion will inevitably be drawn. Aspects for improvement of the situation were identified.

References

1. Melloh M, Staub L, Aghayev E, Zweig T, Barz T, Theis JC, Chavanne A, Grob D, Aebi M, Roeder C (2008) The international spine registry SPINE TANGO: status quo and first results. Eur Spine J 17(9):1201–1209. https://doi.org/10.1007/s00586-008-0665-2

2. Fritzell P, Stromqvist B, Hagg O (2006) A practical approach to spine registers in Europe: the Swedish experience. Eur Spine J 15(Suppl 1):S57–63. https://doi.org/10.1007/s00586-005-1051-y

3. Stromqvist B, Fritzell P, Hagg O, Jonsson B, Sanden B (2013) Swespine: the Swedish spine register: the 2012 report. Eur Spine J 22(4):953–974. https://doi.org/10.1007/s00586-013-2758-9

4. Wood AM, White IR, Thompson SG (2004) Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. Clin Trials 1(4):368–376. https://doi.org/10.1191/1740774504cn032oa

5. Schulz KF, Grimes DA (2002) Sample size slippages in randomised trials: exclusions and the lost and wayward.

Lancet 359(9308):781–785. https://doi.org/10.1016/S0140-6736(02)07882-0

6. McGirt MJ, Parker SL, Asher AL, Norvell D, Sherry N, Devin CJ (2014) Role of prospective registries in defining the value and effectiveness of spine care. Spine (Phila Pa 1976) 39((22 Suppl 1)):S117–128

Affiliations

Bernhard Meyer[1] · Ehab Shiban[1] · Lucia E. Albers[1] · Sandro M. Krieg[1]

Bernhard Meyer
Bernhard.Meyer@tum.de

Ehab Shiban
Ehab.Shiban@tum.de

Lucia E. Albers
Lucia.Albers@tum.de

[1] Department of Neurosurgery, Klinikum rechts der Isar, Technical University Munich, Ismaninger Str. 22, 81675 Munich, Germany