ISQua®

International Journal for Quality in Health Care, 2019, 31(7), G1–G7
doi: 10.1093/intqhc/mzy249
Advance Access Publication Date: 4 January 2019
Methods Article

OXFORD

Methods Article

A novel method to assess data quality in large medical registries and databases

ANDREAS PERREN[1], BERNARD CERUTTI[2], MARK KAUFMANN[3], and HANS ULRICH ROTHEN[4], ON BEHALF OF HE SWISS SOCIETY OF INTENSIVE CARE MEDICINE

[1]Intensive Care Unit, Department of Intensive Care Medicine—Ente Ospedaliero Cantonale, Ospedale Regionale Bellinzona e Valli, 6500 Bellinzona, Switzerland and Faculty of Medicine, University of Geneva, 1206 Geneva, Switzerland, [2]Unit of Development and Research in Medical Education, Faculty of Medicine, University of Geneva, 1206 Geneva, Switzerland, [3]Department of Anaesthesiology, University Hospital, 4031 Basel, Switzerland, and [4]Department of Intensive Care Medicine, Bern University Hospital—Inselspital, 3010 Bern, Switzerland

Address reprint requests to: Andreas Perren, Tel: +0041918119251; Fax: +0041918118022; E-mail: Andreas.Perren@eoc.ch

Editorial Decision 20 November 2018; Accepted 14 December 2018

Abstract

Background: There is no gold standard to assess data quality in large medical registries. Data auditing may be impeded by data protection regulations.
Objective: To explore the applicability and usefulness of funnel plots as a novel tool for data quality control in critical care registries.
Method: The Swiss ICU-Registry from all 77 certified adult Swiss ICUs (2014 and 2015) was subjected to quality assessment (completeness/accuracy). For the analysis of accuracy, a list of logical rules and cross-checks was developed. Type and number of errors (true coding errors or implausible data) were calculated for each ICU, along with noticeable error rates (>mean + 3 SD in the variable's summary measure, or >99.8% CI in the respective funnel-plot).
Results: We investigated 164 415 patient records with 31 items each (37 items: trauma diagnosis). Data completeness was excellent; trauma was the only incomplete item in 1495 of 9871 records (0.1%, 0.0%–0.6% [median, IQR]). In 15 572 patients records (9.5%), we found 3121 coding errors and 31 265 implausible situations; the latter primarily due to non-specific information on patients' provenance/diagnosis or supposed incoherence between diagnosis and treatments. Together, the error rate was 7.6% (5.9%–11%; median, IQR).
Conclusions: The Swiss ICU-Registry is almost complete and data quality seems to be adequate. We propose funnel plots as suitable, easy to implement instrument to assist in quality assurance of such a registry. Based on our analysis, specific feedback to ICUs with special-cause variation is possible and may promote such ICUs to improve the quality of their data.

Key words: medical registry, quality control, accuracy, completeness, funnel-plot

Introduction

Clinical registries are useful tools to monitor the process of patient care, to assist in identifying best practices and to point at possibly inappropriate variation in such processes [1]. In addition, they also constitute an important basis for clinical and epidemiological research [2]. The use of medical registries in Intensive Care Medicine has markedly increased over the last two decades. In Switzerland, participation in collection of such data has become mandatory for all certified Intensive Care Units (ICU) in 2008. The main purpose of the Swiss ICU-Registry (MDSi—Minimal Dataset for ICUs) was to

obtain key indicators of the structure of the single ICUs and of the process of patient care. These indicators are used for the certification of each individual ICU, for accreditation of postgraduate training programs in Intensive Care Medicine for physicians and nurses, for local quality management, for benchmarking and epidemiologic research.

A medical registry is typically developed according to some essential standards: it should be representative of the population that it describes and guarantees both the completeness and accuracy of the collected data [3–5]. Completeness of data is usually ensured on a local level and eventually enforced by the central registry operator. In addition, data accuracy requires a validation concept, e.g. external audits, which are typically based on periodic checks of random samples (whole datasets or just some variables of a defined number of patients from each ICU [6–8]). Extrapolation of these results is then expected to give a general pattern of every single ICU, and eventually of the whole registry. Given the substantial amount of structural and procedural data retrieved every day, the numerous caregivers involved in their collection and the differences and fluctuations over time in education, culture and organization of the various ICUs, some doubts regarding the efficacy of these modest quality control measures might arise [9, 10]. Inaccurate variables seem to be independent of the personal characteristics of data collectors and even repeated audits do not necessarily improve the coding process [10, 11]. Nonetheless, identifying inaccurate data items might be critical, since adjustments in the current data collection procedures could improve data quality (e.g. from manual collection to automatic retrieval of variables from a Patient Data Management System; PDMS) [12, 13].

Because the Swiss laws regarding data protection exclude external checks on local samples without explicit consent from the patients, another method for assessing data quality had to be implemented. The principles of statistical process control have been used to derive funnel plots from clinical datasets. Besides their typical use in measuring the quality of the process of patient care [14–17], funnel plots might also be useful to assess data quality. The aim of the present study was to explore the applicability and usefulness of funnel plots as a novel tool for quality control in critical care registries.

Methods

The Swiss ICU-Registry

The Swiss ICU-Registry (MDSi—Minimal Dataset for ICUs) has been developed according to standards essential for assessing the coverage and accuracy of a clinical database [5]. The aims, contents and required execution (with detailed definitions) of MDSi are well known to all members of the Swiss Society of Intensive Care Medicine (SSICM). The guidelines for MDSi are downloadable as German and French version from SSICMs' website [18].

The MDSi contains variables representing structural characteristics of every ICU and procedural data for every patient admitted to an ICU. Each admission is represented as one patient record, thus a patient admitted to an ICU two or more times during his hospital stay will have a record for each stay in the ICU (the second and any further admission will be labelled as readmission). Variables regarding the structural ICU characteristics are reported once a year and are routinely checked upon the regular process of certification. They include data on number of beds of the ICU, staffing resources and type of accreditation for postgraduate training in Intensive Care Medicine (category of ICU, see Appendix 1). These variables are not part of the research described in this paper. Variables related to the process of care are collected for every patient admitted to an ICU. They are mostly categorical (21 items), some are numerical (5 items) or dates/times (5 items). For trauma patients, the abbreviated injury scale, representing six body regions, is also recorded. Thus, a complete 31-item patient record (37 items for trauma patients) contains pseudonymised information regarding the patient's demographics, diagnostic group (Appendix 2), severity of acute illness (Simplified Acute Physiology Score; SAPS II) [19], intervention prior to ICU admission (Appendix 3), resource use during ICU stay (Nine equivalents of nursing manpower use score; NEMS) for every 8-h nurse shift [20, 21], and outcome. Based on the NEMS and the Sedation Agitation Scale (SAS) [22] or the Richmond Agitation-Sedation Scale (RASS) [23], each patient's eight-hour shift is classified—according to its resource use—into one of four nursing categories (Appendix 4). NEMS and SAPS II are also key components to define the amount of hospital reimbursements in Switzerland (SwissDRG) [24].

Adherence to the MDSi is compulsory for all certified Swiss ICUs, both in public and private hospitals. Each ICU is responsible for data collection, which is generally done manually; just ~30% of the ICUs have implemented some automatic data retrieval (e.g. NEMS, SAPS II) linked to a Patient Data Management System [11, 25]. Collected data are transferred to the central registry database (ProtectData, CH-5623 Boswil; www.protecdata.ch) at least every 6 months. They are routinely checked for completeness according to the defined validation rules. Incomplete data automatically activates the solicitation to the ICU for correction.

During their annual assembly in autumn 2015, the members of the SSICM agreed that all pseudonymised data from the MDSi may be used for research, once an specific request has been approved by the appropriate scientific committee. The Ethics Committee from North-Western Switzerland—corresponding to the legal location of the Swiss Society for Intensive Care Medicine—approved this procedure (EKNZ UBE-15/47).

Quality control

Data of the Swiss ICU-Registry (MDSi) from the years 2014 and 2015 were included in this study. ICUs who exclusively care for neonatological and/or paediatric patients (<16 years) were excluded from the study, since the panel of collected data was slightly diverse. As an example, the system of diagnostic groups and the score used for acute paediatric illness (PIM II) differs from the one used for adults [26].

We checked both, the completeness and the accuracy. Although the central database operator *per se* routinely checks for completeness, we did a further review of all datasets. A patient record was considered complete, if all required items were entered in the appropriate fields. For the analysis of accuracy, a list of logical rules and cross-checks was developed (Table 1). It includes rules related to diagnostics implying some specific treatments, maximal possible range of SAPS II according to the patient's age, length of stay in days equalling number of nurse shifts and correct classification of the patient's severity in Swiss ICU categories conferring to SAS or RASS (Table 1). A data was considered accurate when it was either correct (the five check-rules that revealed clear coding errors in Table 2) or conform to the expected standard (the other analysed data or check-rules). Thus, errors were classified as clear coding errors (e.g. NEMS score >56 points), or as implausible data (e.g. NEMS without basic monitoring). Finally, funnel plots were used to detect ICUs that clearly differ from the average prevalence rates regarding errors or implausible data. For the present study, we defined data as incomplete, if there were missing data.

Table 1 Variables checked for completeness and accuracy

Data completeness	
Patient ID or code >40 characters	
Date of birth	
Gender	
Weight and height, if age <16 years	
Date of ICU admission/discharge	
Outcome upon ICU discharge	
Emergency vs. elective admission	
Readmission <48 h	
Patient's origin before hospital/ICU	
Patient's destination after ICU	
Intervention prior to admission	
Initial diagnosis	
Final diagnosis	
AIS in case of trauma diagnosis	
SAPS II	
NEMS, nine separate variables	
PIM II score	
SAS	
Patient needing microbiologic isolation	
Data accuracy	**Check-rules and expected results**
Age	Date of admission to ICU—date of birth
	Age < 100 years
LOS	Sum of 8 h-shifts/3
	Sum of NEMS-records/3
	LOS < 150 days
	Sum of SSICM categories/3
SAPS II according to age (years)	<40 ≤145 points
	40–59 7–152
	60–69 12–157
	70–74 15–160
	75–79 16–161
	>80 18–163
	SAPS < mean + 3 SD
NEMS	Mean NEMS/shift > nine points*
	Sum of NEMS
	NEMS ≤ 56 points
	NEMS first shift < mean + 3 SD
	NEMS last shift < mean + 3 SD
	No nurse shift without basic monitoring
PIM II score	Age <16 years
	Range 0–100
Trauma as final diagnosis	Presence of ≥1 AIS
Diagnosis–treatment	Otorhinolaringeal neoplasm and surgical intervention
	Cardiac arrest and consecutive ventilation
	ARDS and ventilation (invasive or non-invasive)
	Severe sepsis or septic shock and use of vasopressors
	Cardiogenic shock and use of inotropes
	Trauma diagnostic and acute injury score
General	Sum of SSICM categories <mean + 3 SD
	Sum of SSICM categories >0
	Sum of shifts with SAS > 5 points (RASS > 2) ≤ sum SSICM cat. Ia + Ib + II

AIS, Abbreviated injury scale; LOS, Length of stay; SAPS II, Simplified Acute Physiology Score; NEMS, Nine equivalents of nursing manpower use score; SD, standard deviation; PIM II, Paediatric Index of Mortality; SAS, Sedation Agitation Scale; AIS, Acute Injury Score; ARDS, adult respiratory distress syndrome; SSICM, Swiss Society of Intensive Care Medicine; *Basic monitoring = nine points; SSICM cat, nursing category: see Appendix 3.

Statistical analysis

Every check rule was applied to every patient record. In a second step, the results were aggregated by ICU. Rates of missing data, errors or implausible data were calculated.

We also used funnel plots to identify the ICUs falling outside the 99.8% CI upper band limit, which were defined as outliers [27]. Additional Fisher F tests were performed to investigate whether there was a positive or negative linear trend of the prevalence of

Table 2 Coding errors and implausible data in patient records

	Pat. records (*n*)	Pat. records (%)	ICUs (*n*)
Clear coding errors			
SAPS according to age	1551/164 415	0.94	39
First NEMS > 56 points	2/164 415	0.00	2
Last NEMS > 56 points	3/164 415	0.00	2
Shifts with agitated patient ≤ SSICM category I + II	70/164 415	0.04	5
Trauma diagnostic without AIS	1495/9871	15.15	55
Sub-total	3121/164 415	1.90	65
Implausible data			
Place of origin: 'other'	6478/164 415	3.94	76
Diagnostic group: 'other'	13 807/164 415	8.40	76
Neoplasm of URT without intervention	173/987	17.53	52
Cardiac arrest without mechanical ventilation	486/2569	18.92	73
Cardiogenic shock without vasoactive drugs	2043/5773	35.39	76
ARDS without mechanical ventilation	39/643	6.07	27
Septic shock without vasoactive drugs	1659/6876	24.13	76
SAPS > mean + 3 SD	2304/164 415	1.40	75
Missing sum of NEMS and/or mean NEMS < 9 points	771/164 415	0.47	49
NEMS first shift > mean + 3 SD	415/164 415	0.25	52
NEMS last shift > mean + 3 SD	2977/164 415	1.81	77
Shifts without basic monitoring	113/164 415	0.07	28
Sub-total	12 451/164 415	0.08	77
Total	15 572/164 415	9.47	77

Pat. records, number and percentage of the patient records with coding errors or implausible data; NEMS, Nine equivalents of nursing manpower use score; SSICM, Swiss Society of Intensive Care Medicine; AIS, abbreviated injury score; ARDS, Acute respiratory distress syndrome; URT, upper respiratory tract; SAPS, Simplified Acute Physiology Score; SD, Standard deviation. Mechanical ventilation: includes both invasive and non-invasive ventilation. Vasoactive drugs: includes catecholamines, nitro-vasodilators, phosphodiesterase-3 inhibitors (milrinone), and calcium sensitizer (levosimendan).

cases (errors or implausible data) as a function of the number of patients. A potential link between the rate of clear coding errors or implausible data and the characteristics of the ICU (university hospital ICUs, ICUs offering postgraduate training, ICUs with no postgraduate training) was investigated with Kruskal–Wallis rank sum tests. All analyses were done with R version 3.3.3, R Foundation for Statistical Computing, Vienna, Austria.

Results

A total of 164 415 patient records (covering 1 560 519 nurse shifts) from all 77 certified Swiss ICUs were analysed. The quality control regarding data completeness yielded overall excellent results: trauma diagnosis was the only incompletely registered item, because lacking the abbreviated injury scale (1495/9871; 15.1%). Taking all ICUs together, AIS was missing in 0.1% (0.0%–0.6%; median, 25th–75th percentile) and 10 ICUs had a noticeable error rate.

In 15 572 records (9.5%) from 77 different ICUs, we detected 3121 clear coding errors and 31 265 implausible data (Table 2). Considering the entirety of quality checks (Table 3), the ICUs had a median error rate of 7.6% (5.9%–11%; IQR).

Overall, 19 ICUs showed evidence of clear coding errors; they were largely related to inaccurate scoring of SAPS II (one ICU) and to trauma diagnostic lacking the corresponding abbreviated injury scale (in one ICU, this scale was not used at all). Implausible data were found in patient records from 66 ICUs and were mainly associated to supposedly inaccurate definitions of the patient's diagnostic, their provenance prior to ICU admission, and discrepancies in the logical rules between diagnostics and treatments.

ICUs generating excessively high numbers of inaccurate or implausible data were derived from funnel-plot analysis (Table 3). The presence of outliers was mainly ascribed to the incoherence between reported diagnosis and reported treatments. Figure 1 A is an example for imprecise coding of the diagnostic: In 23 ICUs, 10–61% of the patients were assigned to the unspecific diagnostic group 'other'. Figure 1 B depicts the proportion of patient records without coding of the basic monitoring in the NEMS; five ICUs exceed the 99.8% CI, one of them seem to have managed up to 36 patients without routine check of basic vital parameters. The funnel-plot in Figure 2 gives a summary with all check rules put together: 17 ICUs are flagged outside the 99.8% CI, one ICU presented a coding failure rate as high as 74.5%.

Tertiary ICUs with postgraduate training more precisely coded the diagnostics than other ICUs ($P = 0.0017$). No other significant association was found between coding errors/implausible data and category of ICU (see also Appendix 1).

Discussion

The usefulness of medical registries depends on accuracy and completeness of the collected data [4, 28]. With this study, we show that pre-defined checks combined with statistical methods could be used as a tool for quality control in a large, nation-wide registry of critically ill patients. Accordingly, this allows avoiding the use of random samples of individual records. This is of particular importance, as in Switzerland external checks on local samples without explicit permission by the patients are prohibited by data protection act. Whether the new European Data Protection Regulation, that applies from 25 May 2018, might hamper external audits or peer reviews, remains to be seen [29].

The first favourable result from our research is quasi-completeness of data, an indispensable condition for validity of a registry. In fact, only one out of 32 variables checked was deficient in this respect. This confirms that almost every certified adult Swiss

Table 3 ICUs with unlikely high error rates (Outliers)

	Number of outliers	Error rate per unit (%)		$P_{\text{lin. trend}}$
	n	Outliers min.–max.	All units median (25th–75 percentile)	
Clear coding errors				
SAPS according to age	1	71.6	0.0 (0.0–0.2)	0.78=
First NEMS > 56 points	2	0.01–0.02	0.0 (0.0–0.0)	†
Last NEMS > 56 points	2	0.02–0.03	0.0 (0.0–0.0)	†
Shifts–agitated patient ≤ SSICM category I + II	4	0.2–1.2	0.0 (0.0–0.0)	0.39=
Trauma diagnostic without AIS	10	14.0–99.8	0.1 (0.0–0.6)	0.037 ↑
Implausible data				
Imprecisely defined place of origin	21	5.4–17.1	3.8 (1.9–5.7)	0.14=
Imprecisely defined diagnostic	23	10.3–61.3	7.6 (4.5–10.8)	0.062=
Neoplasm of URT without intervention	4	36.4–83.9	27.7 (3.8–57.0)	0.006 ↓
Cardiac arrest without mechanical ventilation	5	60.0–100.0	28.6 (16.7–40.0)	<0.0001 ↓
Cardiogenic shock without vasocative drugs	16	54.4–81.8	35.8 (26.8–53.3)	0.019 ↓
ARDS without ventilation	5	25.0–100.0	0.0 (0.0–12.5)	0.12=
Septic shock without vasoactive drugs	13	34.0–74.1	23.4 (16.7–34.0)	0.032 ↓
SAPS > mean + 3 SD	12	2.2–5.3	1.2 (0.6–1.9)	0.38=
Missing sum of NEMS and mean NEMS < 9 points	7	2.9–11.3	0.1 (0.0–0.2)	0.17=
NEMS, first shift > mean + 3 SD	7	0.5–1.5	0.1 (0.0–0.2)	0.015 ↑
NEMS, last shift > mean + 3 SD	10	2.8–9.3	1.4 (0.8–1.8)	0.13=
Shifts without basic monitoring	5	0.4–2.1	0.0 (0.0–0.1)	0.22=
All checks together	17	11.3–74.5	7.6 (5.9–11.0)	0.554=

Outliers, range of patient records (%) with erroneous/implausible data of ICUs exceeding the 99.8% CI of the funnel plots or mean > 3 SD; URT, upper respiratory tract; AIS, acute injury score; SAPS, Simplified Acute Physiology Score; NEMS, Nine equivalents of nursing manpower use score; SSICM, Swiss Society of Intensive Care Medicine.

$P_{\text{lin. trend}}$: *P*-value Fisher's test investigating the presence of a linear trend regarding an error rate that depends on the number of patients; †: rate of error too small; =, no significant trend; ↓ (↑), decreasing (increasing) trend along with higher number of patients. Mechanical ventilation: includes both invasive and non-invasive ventilation. Vasoactive drugs: includes catecholamines, nitro-vasodilators, phosphodiesterase-3 inhibitors (milrinone), and calcium sensitizer (levosimendan).

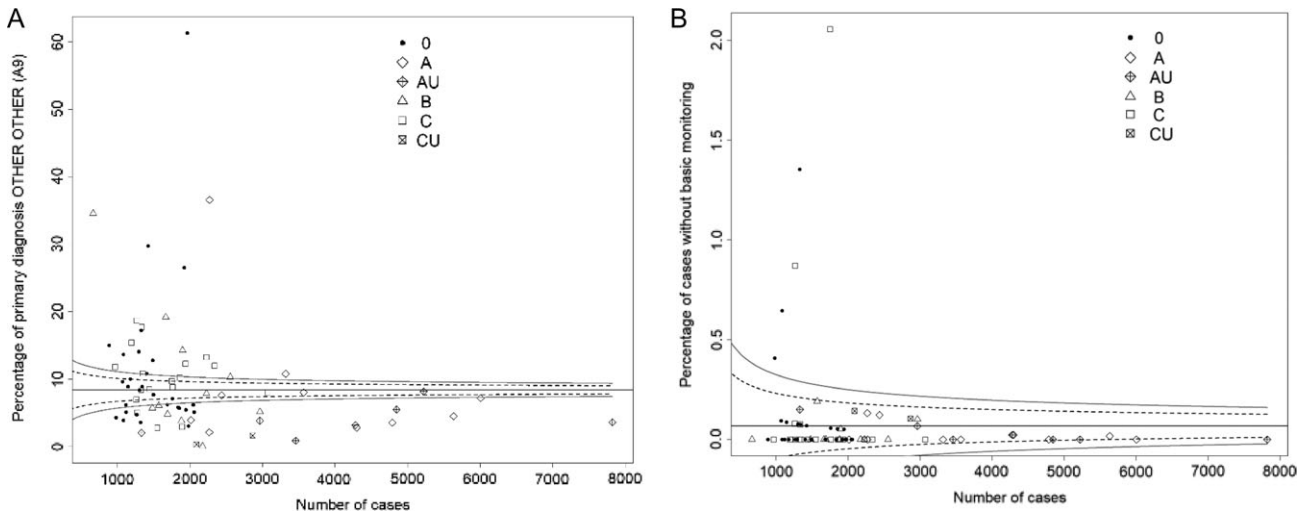

Figure 1 A. Funnel-plot depicting the frequency of diagnosis 'other' vs. number of patients: 23 ICUs are flagged above the 99.8% CI upper limit. B. Funnel-plot illustrating the proportion of patients apparently managed without basic monitoring. Five ICUs are flagged above the 99.8% CI upper limit. Category of ICU: *0*, ICUs without postgraduate training; *A*, (*B*, *C*) ICUs with 36 (18, 12) months of postgraduate training; *U*, tertiary (universitary) ICU, always combined with a specific duration (A–C).

ICU provides all data for each patient necessary to calculate key process indicators. This is an excellent result as the portion of missing data in medical registries may be substantial, reaching up to 17% [7, 30–32]. In MDSi, the high level of completeness of data is assured by an automatic data completeness-check by the central registry operator. Obviously, this check is missing for one single variable and this shortcoming will be soon corrected.

Data accuracy was ascertained by means of 31 check rules developed by the authors of this study. Clear coding errors, i.e. logically not plausible combinations of data, were frequently related to

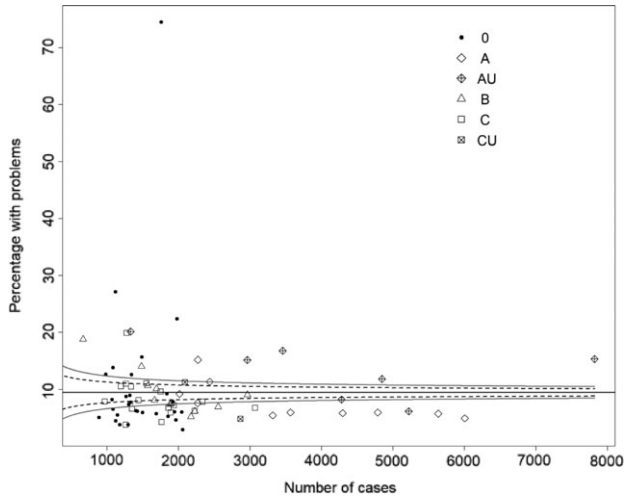

Figure 2 Funnel-plot depicting the global proportion of erroneous or implausible coding (all check-rules analysed together): 17 ICUs are flagged above the 99.8% CI upper limit. Category of ICU: *0*, ICUs without postgraduate training; *A*, (*B*, *C*) ICUs with 36 (18, 12) months of postgraduate training; *U*, tertiary (universitary) ICU, always combined with a specific duration (A–C).

inadequate scoring of NEMS and SAPS II. Indeed, this problem is well known and has been described previously [11, 25]. Finally, to check for implausible data, we searched for extreme outliers either by looking at normality distributions or by means of funnel plots. Of note, almost every single ICU had some records with implausible data. Of course, a median error rate of 7.6% might appear rather high at first glance. However, as our method of quality control is completely novel, our results cannot be easily compared.

As already noted by Seneca *errare humanum est* meaning that a certain error rate is unavoidable and that there is hardly any measure to prevent it. Such error may result in variability in measures of patient care and in performance indicators [27]. Typically, it is assumed that variation in patient care and in patient-mix in an otherwise stable process are key contributors to such variability. This kind of variation, a common-cause failure (i.e. inherent in every process, endogenous), is expected to be found within the 99.8% confidence interval (CI) of a funnel-plot [33]. On the other hand, outliers —special-cause failure (i.e. exogenous, an intermittent and unpredictable process that arises from unusual circumstances)—deserve special attention. As recently mentioned by Verbrug *et al.* [34], we argue that besides unmeasured covariates, problems in case-mix correction and policy choices, errors in registration or coding are important contributing factors. Once identified, correcting the coding procedure should be a feasable issue in most instances. Thus, funnel plots could also allow for rapid visual comparison of coding quality. Units having data points outside the control limits will need to appraise their local practices to avoid future errors in coding.

In fact, there is a second half to the famous line of the roman philosopher, i.e. *sed perseverare diabolicum* (to persist in error is diabolical). Such quality control, warranting feedback to concerned ICUs and subsequent action in their coding practice, may lay the foundations for continuous quality assurance in a registry. In this sense, the SSICM may consider introducing yearly quality controls. This would be particularly useful as our MDSi actually gives a complete picture of the whole nation's critical care activity, insofar as all certified Swiss ICUs are obliged to furnish these key process indicators.

Considering the continuous improvement in our clinical and administrative tasks as a very important matter, detecting specific flaws in the documentation of our work is essential. Like other analysis methods, funnel plots derived from a statistical process control, might produce both false negative and positive results. In this sense, lying within or outside the control limits does not necessarily imply that the corresponding ICU performs well or bad. This may be particularly true for the logical rules between diagnosis and treatment. Let's take, as an example, the logical rule 'cardiac arrest and mechanical ventilation'. Indeed, the literature reports that about 34% of in-hospital cardiac arrests are managed without intubation and mechanical ventilation [35]. In our registry, flagged ICUs (>99.8% CI) in the corresponding funnel-plot had rates above 60%, suggesting either a problem in coding of diagnosis or treatment, or the possibility of non-standard clinical management. Thus, funnel plots may be powerful instruments, capable to detect possible special-cause variation that needs investigation. Likewise, they might avert unsuitable exploration of an outcome resulting from common-cause variation.

There are strengths and limitations to this approach of quality control in a registry. Compared to the typical method based on external audits [4, 10] or checks on random samples [36, 37], various differences stand out. Checking for special-cause variation is rather easy to perform. In spite of analysing numerous data in few patient records by external audit, a check of pre-specified data in all records can be done repeatedly and without use of large resources. There are however also some limitations. As discussed previously, it has to be kept in mind that special-cause variation can also point at true—and at times clinically relevant—differences in the process of care, but also at differences in case-mix, at inappropriate models of risk adjustment, or at the statistical model itself used to estimate the degree of over-dispersion [38]. However, once detected as outliers, some data may be difficult to verify without going back to the patient records. Examples are the diagnosis, the patient's provenance or his destination after ICU stay. Similarly, a point-to-point check of the various items within the SAPS II and NEMS scores will be extremely laborious, if not impossible.

In conclusion, the Swiss ICU-Registry (MDSi) is next to complete and data quality seems to be adequate. Funnel plots might be suitable instruments for quality control of medical registries. Based on our analysis, specific feedback to ICUs with special-cause variation is possible and might assist such ICUs to improve the quality of their data.

Supplementary material

Supplementary material is available at *International Journal for Quality in Health Care* online.

Funding

This study was supported by a grant from the Research Fund of the Swiss Society of Intensive Care Medicine.

Contributors

Each author has contributed significantly to the manuscript and meets the requirements of authorship as specified by the Uniform Requirements for Manuscripts Submitted to Biomedical Journals.

Patient consent

The Ethics Committee from North-Western Switzerland—corresponding to the legal location of the Swiss Society for Intensive Care Medicine—approved this procedure (EKNZ UBE-15/47).

References

1. Md Emdadul Hoque D, Ruseckaite R, Lorgelly P *et al*. Cross-sectional study of characteristics of clinical registries in Australia: a resource for clinicians and policy makers. *Int J Qual Health Care* 2018;**30**:192–9.
2. Riddersholm S, Christensen S, Kragholm K *et al*. Organ support therapy in the intensive care unit and return to work: a nationwide, register-based cohort study. *Intensive Care Med* 2018;**44**:418–27.
3. Black N. Developing high quality clinical database. *BMJ* 1997;**315**:381–2.
4. Arts DGT, De Keizer N, 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. https://doi.org/10.1197/jamia.M1087.
5. Black N, Payne M, on behalf of the DoCDat Development Group. Directory of Clinical Databases: improving and promoting their use. *Qual Saf Health Care* 2003;**12**:348–52.
6. Fisher ES, Whaley FS, Krushat WM *et al*. The accuracy of Medicare's hospital claims data: progress has been made, but problems remain. *Am J Public Health* 1992;**82**:243–248.
7. Kelly M, Lamah M. Evaluating the accuracy of data entry in a regional colorectal cancer database: implications for national audit. *Colorectal Dis* 2007;**9**:337–339.
8. Stow PJ, Hart GK, Higlett T *et al*. ANZICS Database Management Committee. Development and implementation of a high-quality clinical database: the Australian and New Zealand Intensive Care Society Adult Patient Database. *J Crit Care* 2006;**21**:133–41.
9. Hewson-Conroy KM, Tierney LT, Burrell AR. Assessment and perceptions of intensive care data quality, reporting and use: a survey of ICU directors. *Anaesth Intensive Care* 2012;**40**:675–82.
10. Willis CD, Jolley DJ, McNeil JJ *et al*. Identifying and improving unreliable items in registries through data auditing. *Int J Qual Health Care* 2011;**23**:317–23. doi:10.1093/intqhc/mzr004.
11. Previsdomini M, Cerutti B, Merlani P *et al*. SwissScoring—a nationwide survey of SAPS II assessing practices and its accuracy. *Swiss Med Wkly* 2014;**144**:w14090.
12. Arts D, de Keizer N, Scheffer G-J *et al*. Quality of data collected for severity of illness scores in the Dutch National Intensive Care Evaluation (NICE) registry. *Intensive Care Med* 2002;**28**:656–9.
13. Bosman RJ, Oudemane van Straaten HM, Zandstra DF. The use of intensive care information systems alters outcome prediction. *Intensive Care Med* 1998;**24**:953–8.
14. Gale CP, Roberts AP, Batin PD *et al*. Funnel plots, performance variation and the Myocardial Infarction National Audit Project 2003–2004. *BMC Cardiovasc Disord* 2006;**6**:34.
15. Kunadian B, Dunning J, Roberts AP *et al*. Funnel plots for comparing performance of PCI performing hospitals and cardiologists: demonstration of utility using the New York hospital mortality data. *Catheter Cardiovasc Interv* 2009;**73**:589–94. doi:10.1002/ccd.21893.
16. Dover DC, Schopflocher DP. Using funnel plots in public health surveillance. *Popul Health Metr* 2011;**9**:58. doi:10.1186/1478-7954-9-58.
17. TARN. The Trauma Audit & Research Network. Clinical report issue 2-July 2017. I: Core measures for all patients, II: Orthopaedic injuries 2017. Accessed September 26, 2018 at: https://www.tarn.ac.uk/content/downloads/53/July%2017%20Revised%20Clinical%20report%20-%20TU%20example.pdf
18. Minimaler Datensatz der SGI – MDSi; May 11, 2015; German edition: https://www.sgi-ssmi.ch/de/datensatz.html. Accessed [December 10, 2016].
19. Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicentre study. *JAMA* 1993;**270**:2957–63.
20. Reis Miranda D, Moreno R, Iapichino G. Nine equivalents of nursing manpower use score (NEMS). *Intensive Care Med* 1997;**23**:760–5.
21. Rothen HU, Küng V, Ryser DH *et al*. Validation of 'nine equivalents of nursing manpower use score' on an independent data sample. *Intensive Care Med* 1999;**25**:606–11.
22. Riker RR, Picard JT, Fraser GL. Prospective evaluation of the Sedation-Agitation Scale for adult critically ill patients. *Crit Care Med* 1999;**27**:1325–9.
23. Sessler CN, Gosnell MS, Grap MJ *et al*. The Richmond Agitation-Sedation Scale: validity and reliability in adult intensive care unit patients. *Am J Respir Crit Care Med* 2002;**166**:1338–44.
24. SwissDRG; German edition: http://www.www.swissdrg.org/de. (2018) Accessed [May 4, 2018].
25. Perren A, Cerutti B, Merlani P *et al*. SwissScoring—a nationwide survey of NEMS assessing practices and its accuracy. *Acta Anaesthesiol Scand* 2014;**58**:478–86. doi:10.1111/aas.12283.
26. Slater A, Shann F, Pearson G. Paediatric Index of Mortality (PIM) Study Group. PIM2: a revised version of the Paediatric Index of Mortality. *Intensive Care Med* 2003;**29**:278–85.
27. Spiegelhalter DJ. Funnel plots for comparing institutional performance. *Stat Med* 2005;**24**:1185–1202.
28. Nelson EC, Dixon-Woods M, Batalden PB *et al*. Patient focused registries can improve health, care, and science. *BMJ* 2016;**354**:i3319.
29. Data Protection in the European Union: https://ec.europa.eu/info/law/law-topic/data-protection_en. Accessed [June 08, 2018].
30. Beretta L, Aldrovandi V, Grandi E *et al*. Improving the quality of data entry in a low-budget head injury database. *Acta Neurochir* 2007;**149**:903–9. Doi;10.1007/s00701-007-1257-3.
31. McClean K, Mullany D, Huckson S *et al*. Identification and assessment of potentially high-mortality intensive care units using the ANZICS Centre for Outcome and Resource Evaluation clinical registry. *Crit Care Resusc* 2017;**19**:230–8.
32. Christiansen CF, Møller MH, Nielsen H *et al*. The Danish intensive care database. *Clin Epidemiol* 2016;**8**:525–30.
33. Mayer EK, Bottle A, Rao C *et al*. Funnel plots and their emerging application in surgery. *Ann Surg* 2009;**249**:376–83.
34. Verburg IW, Holman R, Peek N *et al*. Guidelines on constructing funnel plots for quality indicators: a case study on mortality in intensive care unit patients. *Stat Methods Med Res* 2017;**1**:962280217700169. doi:10.1177/0962280217700169.
35. Andersen LW, Granfeldt A, Callaway CW *et al*. American Heart Association's Get With The Guidelines–Resuscitation Investigators. Association between tracheal intubation during adult in-hospital cardiac arrest and survival. *JAMA* 2017;**317**:494–506. doi:10.1001/jama.2016.20165.
36. Horton EE, Krijnen P, Molenaar HM *et al*. Trauma West Research Group. Are the registry data reliable? An audit of a regional trauma registry in the Netherlands. *Int J Qual Health Care* 2017;**29**:98–103.
37. Schmidt M, Schmidt SA, Sandegaard JL *et al*. The Danish National Patient Registry: a review of content, data quality, and research potential. Clinical. *Epidemiol* 2015;**7**:449–90.
38. Spiegelhalter DJ. Handling over-dispersion of performance indicators. *Qual Saf Health Care* 2005;**14**:347–51.