World J Surg (2018) 42:54–60
DOI 10.1007/s00268-017-4172-7

ORIGINAL SCIENTIFIC REPORT

Validation of an Electronic Surgical Outcomes Database at Mbarara Regional Referral Hospital, Uganda

Geoffrey A. Anderson[1] · Lenka Ilcisin[2] · Joseph Ngonzi[3] · Stephen Ttendo[4] · Deus Twesigye[5] · Noralis Portal Benitez[6] · Paul Firth[7] · Deepika Nehra[2]

Published online: 7 August 2017

Abstract

Background Accurate, complete and sustainable methods of tracking patients and outcomes in low-resource settings are imperative as we launch efforts to improve surgical care globally. The Surgical services QUality Assessment Database (SQUAD) at the Mbarara Regional Referral Hospital in Uganda is one of very few electronic surgical databases in a low-resource setting. We evaluated the completeness and accuracy of SQUAD.

Methods Data were prospectively collected on 20 of the most clinically relevant variables captured by SQUAD for all general surgery patients admitted to MRRH over a two-week period. Patients were followed until discharge, death or hospital day 30; whichever occurred first. These data were compared to that in SQUAD for the same time period for completeness and accuracy.

Results Of 186 unique patients seen over the two-week period, 172 (92.5%) were captured by SQUAD. The capture rate was greater than 86% for each of the 20 variables evaluated, except American Society of Anesthesiologists score, which had a 69% capture rate. Regarding accuracy, there was almost perfect agreement for 16/20 variables (all $k > 0.81$), substantial agreement for 2/20 variables (k 0.63, 0.73) and moderate agreement for the remaining 2/20 variables (k 0.43, 0.48) between SQUAD and the prospectively collected data.

Conclusion SQUAD is an electronic surgical database that has been implemented and sustained in a low-resource setting. For the 20 variables evaluated, the data within SQUAD are highly complete and accurate. This database may serve as a model for the development of additional surgical databases in low-resource environments.

Geoffrey A. Anderson and Lenka Ilcisin are Co-first authors.

✉ Deepika Nehra
dnehra@bwh.harvard.edu

[1] Department of Surgery, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02144, USA

[2] Department of Surgery, Brigham and Women's Hospital, 75 Francis St, ASB II L1, Boston, MA 02115, USA

[3] Department of Obstetrics and Gynecology, Mbarara University of Science and Technology, PO Box 1410, Mbarara, Uganda

[4] Department of Anesthesia and Critical Care, Mbarara University of Science and Technology, PO Box 1410, Mbarara, Uganda

[5] Department of Surgery, Mbarara Regional Referral Hospital, Mbarara, Uganda

[6] Department of Surgery, Mbarara University of Science and Technology, PO Box 1410, Mbarara, Uganda

[7] Department of Anesthesia, Critical Care and Pain Management, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02144, USA

Introduction

In 2015, the Lancet Commission on Global Surgery brought to light several key findings that demonstrated the staggering human and economic consequences of untreated surgical conditions in low- and middle-income countries (LMICs) and urged the development of broad-based health-systems solutions [1]. One such solution was the development of surgical outcomes databases that would facilitate the understanding of the current disease burden and outcomes. In high-income countries, accurate and detailed databases exist. One of the most well-known examples of such a database is the National Surgical Quality Improvement Project (NSQIP) [2, 3]. This database tracks over 130 variables on thousands of patients across hundreds of hospitals throughout the USA. By providing accurate and timely clinical data, NSQIP has brought to light the powerful effect that such a database can have on helping both hospitals and providers achieve safer surgery and better patient care [4].

The real question becomes: is this sort of a database, and the resultant safer surgery and better patient care, something we can strive for in LMICs? Currently, there are very few well-established validated electronic surgical registries in low-income countries (LICs). Data collection in most LICs is done entirely through handwritten logbooks and paper charts [5–8]. This method of record keeping presents an enormous barrier to using this data for any purpose, whether it be patient care, policy work, resource allocation, research or quality improvement initiatives [5–9]. As a result, very few LMICs are aware of the true burden of surgical disease and the associated outcomes in their population. This stifles initiatives to improve access to safe surgical care, the very mission as stated by the Lancet Commission on Global Surgery.

To address this issue, efforts are underway to develop sustainable and effective electronic surgical outcomes databases appropriate for the limited resource environment. One of these is the Surgical services QUality Assurance Database (SQUAD) that was developed in Uganda through a partnership between Mbarara University of Science and Technology (MUST) and its affiliated hospital Mbarara Regional Referral Hospital (MRRH) and the Massachusetts General Hospital (MGH). SQUAD was initiated in 2013; and the development and early success of this database was described as a teaching case during the Lancet Commission on Global Surgery [10]. Although this database exists and may be a powerful example of what an electronic database can provide for LMICs, no assessment of data quality has been performed to date. The purpose of this study is to evaluate the completeness and validity of SQUAD.

Methods

MRRH hospital and record system

MRRH is a 600-bed, government referral hospital in southwest Uganda that serves a catchment area of over 3 million people and is the specialty referral center for a region of 8 million [11, 12]. It has four operating theaters and a number of anesthesiologists and surgeons, including subspecialists. All patients are admitted via the emergency department where a paper chart is created that stays with the patient for the duration of their hospital stay. With rare exception, a new chart is created for each patient encounter, even if the same patient has been admitted previously. Patients are also tracked in various logbooks throughout the hospital which are maintained by nursing and surgical staff.

SQUAD database

Data entry into SQUAD was initiated in 2013, and SQUAD currently enrolls all patients admitted to the surgical service. OpenMRS was used to create the database, and all data are stored on an encrypted local network within the hospital. OpenMRS is an open-source electronic medical record system designed for use in low-resource settings [13]. Two data clerks are responsible for data entry, and there is an onsite database manager and statistician. The database is overseen by a team of physicians from a variety of specialties including surgery, anesthesia and obstetrics and gynecology.

Patients are admitted to the surgical services via the accident and emergency ward, where the admission is noted in a logbook and a patient file is created. The paper chart accompanies the patient throughout their hospital stay. After patients are discharged, SQUAD data clerks collect the patient charts from each surgical ward and manually enter the data into the electronic database. The paper charts are then sent to medical records for filing and storage. Patient encounters are also captured from the ward and operating room logbooks, in order to capture patients whose charts are misplaced.

Each patient encounter receives a unique SQUAD identifier based on the chart, and these are linked with the patient's name, age and address. As charts are rarely reused across multiple admissions, demographic data are used to identify possible duplicate patients and to link multiple encounters. Over a hundred variables can be captured in the database that broadly cover demographic information, admission data, procedure data (both operative and anesthetic), and disposition. Additional details regarding traumatic injuries, oncologic diagnosis and pregnancy outcomes are recorded where relevant.

Table 1 Variables captured during the validation of SQUAD

Variable	Format
Name	Free text
Medical record number	Free text
Age	Continuous (nearest year)
Gender	Dichotomous (male vs female)
Date of admission	Day/month/year
Admission diagnosis	Free text
Operation	Dichotomous (yes/no)
Type of operation	Free text
Date of operation	Day/month/year
Anesthesia type	General, regional, local, combination, other
Surgeon	Free text
Anesthesiologist	Free text
ASA class	Ordinal (1–5)
Urgency of operation	Dichotomous (emergent vs elective)
ICU admission	Dichotomous (yes/no)
Mechanical ventilation	Dichotomous (yes/no)
Discharge diagnosis	Free text
Complication	Dichotomous (yes/no)
Disposition	Dichotomous (alive or dead at departure from hospital)
Date of disposition	Day/month/year

Further details regarding SQUAD have been described previously in the form of a teaching case for the Lancet Commission on Global Surgery [10] (This two part teaching case can be found at http://www.lancetglobalsurgery.org/teaching-cases, Part A: http://docs.wixstatic.com/ugd/346076_40106c3b9bda42a2854fbc0cf8d1614e.pdf, Part B: http://docs.wixstatic.com/ugd/346076_bacf80f10dc246ff81157a46b04787cf.pdf).

Power calculation

A power analysis was performed to estimate sample size for a study with 80% power to determine a 5% difference in completeness of patient capture with an alpha level of 0.05. Based on database entries from 2014, a two-week period of enrollment would capture the 150 patients required for a sufficient sample size. Ethical approval was obtained from the Institutional Review Committee at MUST, the Ugandan National Committee for Science and Technology (UNCST) and from the Institutional Review Board at Boston Children's Hospital.

Prospective data collection

Prospective data were collected for all patients admitted to the surgical services at MRRH over a two-week period in November 2015. Otolaryngology patients were excluded because that component is being validated separately.

Twenty variables, chosen on the basis of a review of surgical outcomes literature to determine the variables most important in quality improvement and outcomes research [14–16], were captured (Table 1). We ensured that the chosen variables allow for calculation of important metrics recommended by the World Bank and the Lancet Commission on Global Surgery, such as surgical volume and post-operative mortality rate [1, 17]. Of note, the variable "complication" simply denotes whether any complication (specifically surgical site infection, wound dehiscence or deep venous thrombosis) was recorded in the chart.

Prospective data collection was completed by direct observation. Data collectors attended morning rounds in the emergency department and all surgical wards, in addition to performing direct observation in the operating theaters and intensive care unit (ICU). No direct observation occurred after dark due to safety concerns. Each morning, data collectors met with overnight staff to complete data collection on patients admitted overnight. Data collection continued until discharge, hospital day 30 or death, whichever occurred first. Most of the variables collected allowed for simple objective observations (e.g., date of operation, surgeon, gender). For those variables that had a subjective component [e.g., diagnosis, operation, American Society of Anesthesiologist Score (ASA)], we observed what the clinician recorded in the logbooks and paper charts.

Completeness and accuracy

The prospectively collected data were compared to the data entered into SQUAD over the same time period for completeness and accuracy.

The completeness of the SQUAD database was defined by the proportion of all patients and variables included in the prospective data captured from all data collection methods that were captured by SQUAD.

Accuracy of data within the SQUAD database was assessed by comparing data points between the SQUAD database and the prospectively collected data for the 164 patients represented in both cohorts. Accuracy was assessed in two ways. We first determined if the data collected in SQUAD agreed with that collected prospectively. Two individuals independently rated each variable for every patient as "agree" or "disagree" between the 2 data collection methods. When the two raters disagreed, an

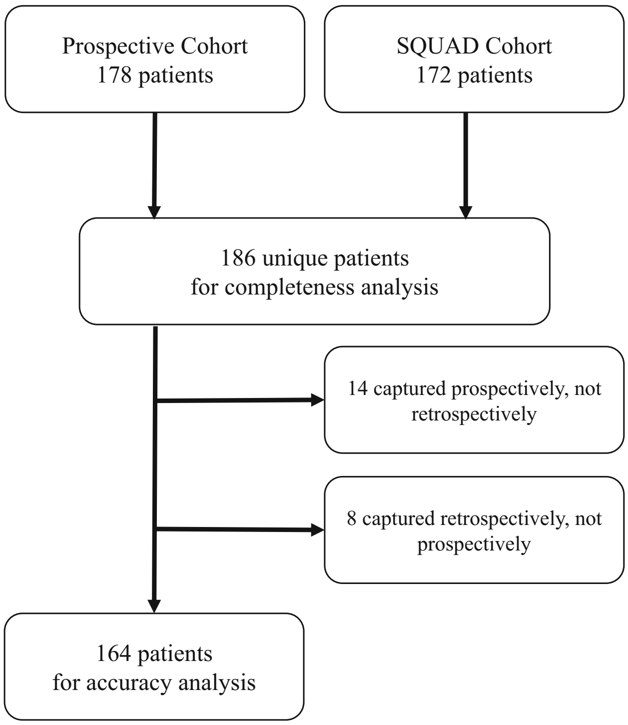

Fig. 1 Patient flow chart. Number of patients in each data collection and comparison group

arbitration was performed with a third rater until consensus was reached. Because time of admission is not captured in the logbook or charts, dates of admission were considered to be concordant if the two dates were within one day (to exclude admissions around midnight). Because the actual date of discharge is not recorded on weekends, date of discharge was considered concordant if the two dates were within 3 days.

Inter-rater reliability between prospective data collection and SQUAD was determined by calculation of a kappa statistic for each variable. This is a more sensitive measure for low-frequency observations, because it takes into account the percentage of matches that would happen by chance. We used the standard qualitative descriptive terms associated with a range of kappa values (0.01–0.2 no to slight agreement, 0.21–0.4 fair, 0.41–0.60 moderate, 0.61–0.80 substantial, 0.81–1.0 excellent or almost perfect agreement) [18].

All statistical analysis was performed using STATA 14 (College Station, TX).

Results

During the two-week period of prospective data collection, 178 patients were captured. Over the same period, 172 patients had encounters recorded in the SQUAD database. Fourteen patients captured prospectively were not captured in SQUAD, and SQUAD captured eight patients not captured in the prospective data collection (Fig. 1), for a total of 186 patients.

Basic demographic characteristics between the SQUAD database cohort and the prospectively collected cohort were very similar (Table 2). Male patients accounted for 76%, with an average age of 27 years and a range between 0 and 98 years. Just over one-third of all patients admitted to the surgical services underwent at least one operative procedure, and approximately half of these procedures were classified as emergent. Median ASA was 2, and less than one-third of patients had an ASA of $\geq$3. Only 5% of patients were admitted to the ICU at any point during their hospital course, and the postoperative complications captured (surgical site infection, wound dehiscence, deep venous thrombosis) were rare. Four patients were still in the hospital at the 30-day point and were thus censored. The median hospital length of stay was approximately 3 days, with an interquartile range of 1–7 days. Thirteen patients in each cohort died, producing an overall mortality of approximately 7–8%.

Overall, SQUAD was complete for the variables of interest. As seen in Table 3, all variables except ASA had a data point capture rate of greater than 85%. ASA was recorded for 69% of patients who had an ASA collected during the prospective data collection. The 14 patients not identified in SQUAD, trended toward younger ages. None of these patients were admitted to the ICU or had emergent surgeries.

A comparison of the accuracy of data is also displayed in Table 3. Sixteen of the 20 variables were found to be more than 90% accurate. Three of the variables (anesthesia type, operation urgency and surgeon) were accurately recorded between 80 and 90% of the time in SQUAD. Finally, ASA had only 54% concordance with prospective data collection.

As expected from the percent concordance, 16 of the 20 variables had a kappa statistic of >0.80, or “almost perfect agreement”. Two variables (procedure urgency and anesthesia type) had a kappa statistic between 0.6 and 0.8 or “substantial agreement”. The last two variables (complication and ASA) had a kappa statistic of 0.48 and 0.43, respectively, or “moderate agreement”.

Discussion

We performed a prospective study to determine the validity a surgical outcomes database (SQUAD) in a referral hospital in Uganda. Overall, patient capture in SQUAD was excellent. Of 186 surgical patients seen over the study period, 172 (92.5%) were captured by SQUAD. The

Table 2 Descriptive statistics of the different cohorts

	Prospective cohort ($N = 178$)	SQUAD cohort ($N = 172$)	Patients missed in SQUAD ($N = 14$)
Male sex	135 (76)	131 (76)	10 (71)
Age (years)	23 (7–37)	24 (8–37)	13.5 (4–30)
Range	0–98	0–98	0–78
<1	14 (7)	12 (6)	2 (14)
1–18	59 (31)	55 (29)	7 (50)
19–64	89 (47)	90 (48)	4 (29)
>65	26 (14)	31 (16)	1 (7)
Surgical procedure performed[a]	70 (38)	65 (37)	4 (29)
Emergent Procedure	41 (59)	32 (49)	0 (0)
ASA Score	2 (1–3)	2 (1–3)	2.5 (1.5–3)
1	24 (36)	13 (25)	1 (25)
2	23 (34)	25 (49)	1 (25)
3	14 (21)	9 (18)	0
4	4 (6)	2 (4)	0
5	2 (3)	2 (4)	0
ICU admission	8 (4)	8 (5)	0
Complication	8 (4)	4 (2)	1 (7)
Hospital length of stay (days)	3 (1–7)	3 (1–7)	2.5 (1–7)
Death	13 (7)	13 (8)	1 (7)

Comparison of basic demographic and basic inpatient characteristics between Prospective and SQUAD cohorts and those patients missing from SQUAD

All data presented as number of patients (%) or median (IQR)

[a] Two patients had two procedures, total number of procedures is reported, (%) represents the percent of patients having one or more surgical procedures

individual variable capture rate in SQUAD was more than 85% for all variables examined with the exception of ASA. ASA was recorded in SQUAD 69% of the time. SQUAD was also highly accurate. We found 16 of the 20 variables were accurately recorded more than 90% of the time. Another three—type of anesthesia, operation urgency and surgeon—were 80–90% accurate. Finally, ASA was accurately recorded only 54% of the time. The inter-rater reliability for 16 of the 20 variables had near perfect agreement (k 0.8–1.0). Operation urgency and anesthesia type had substantial agreement (k 0.6–0.8). ASA and complications had only moderate agreement (k 0.4–0.6).

These data suggest that SQUAD is a valid database for the selected variables. Nearly all patients admitted to the surgical services during the study period were included in SQUAD. Important variables such as age, gender, dates of admission and discharge, diagnosis, operation, and ultimate disposition are highly accurate. This will allow for calculations of important outcomes metrics and for basic risk adjustment.

ASA has been shown to be one of the most important variables for surgical risk stratification [15, 16]. Our data suggest that there is room for improvement in ASA capture by SQUAD before it can be used with confidence. Our data collection highlighted some ways that this might be improved, especially with regard to logbook review. If you consider ASA accuracy as within ± 1 of that recorded during prospective data collection, the accuracy increases from 54 to 67%. Additionally, if you consider only patients for whom SQUAD captured an ASA and classify ASA as accurate within ± 1, the accuracy increases to 91%.

An outcome that is frequently reported in the surgical literature is postoperative complication rate. In our study, the variable “complication” was found to have a k of 0.48, the lowest of all the variables. We found during our data collection that due to resource constraints common surgical complications included in other outcomes databases, such as thromboembolic, are rarely diagnosed or documented. We do not feel, therefore, that any type of complication, or even the overall rate of postoperative complications, can be reliably assessed with SQUAD.

There are a number of other variables that are often recorded in surgical outcomes research, such as imaging and laboratory values, that we did not assess. Albumin, in particular, has been shown to be an important variable in risk stratification in NSQIP [15, 16]. We did not even

Table 3 SQUAD Database Completeness and Accuracy

	Completeness (% complete)	Accuracy (% inter-rater agreement)	Kappa statistic (inter-rater reliability (CI))
Name	92	100	1.00 (0.99–1.01)
MRN	92	90	0.88 (0.87–0.89)
Age	92	96	0.96 (0.94–0.99)
Sex	92	99	0.98 (0.83–1.14)
Admission date	92	98	0.97 (0.93–1.02)
Admission diagnosis	92	97	0.97 (0.94–1.00)
Disposition date	90	93	0.92 (0.89–0.96)
Disposition	90	99	1.00 (0.84–1.15)
Discharge diagnosis	89	96	0.94 (0.92–0.97)
ICU admission	91	98	0.87 (0.72–1.01)
Ventilator use	91	99	0.85 (0.72–0.99)
Complication	91	97	0.48 (0.34–0.62)
Surgery performed	92	96	0.93 (0.78–1.09)
Procedure date	88	91	0.92 (0.85–0.98)
Anesthesia type	88	80	0.63 (0.44–0.82)
ASA	69	54	0.43 (0.32–0.55)
Operation urgency	88	85	0.73 (0.51–0.95)
Procedure description	88	91	0.94 (0.89–0.98)
Surgeon	86	82	0.82 (0.74–0.91)
Anesthesiologist	88	95	0.89 (0.79–0.98)

The kappa statistic is a measure of inter-rater reliability and can be interpreted as follows: 0.01–0.2: no to slight agreement, 0.21–0.4: fair agreement, 0.41–0.60: moderate agreement, 0.61–0.80: substantial agreement, 0.81–1.0: excellent or almost perfect agreement [17]

attempt to measure these variables because they are almost never collected or recorded in the charts at MRRH. In many LICs, imaging and laboratory investigations are seldom used because they are either not readily available, or only available at great cost to the patient. We cannot recommend that SQUAD be used to examine outcomes that rely on the use of imaging or laboratory investigations.

There are several limitations to this study. A comprehensive assessment of database validity looks at six different factors: completeness, accuracy, precision, correctness, consistency and timeliness [19, 20]. We were able to directly assess only two of these six variables in this study; but, arguably, these are the most important two variables. It is not unusual for a database validation to assess only some of the aforementioned parameters. It has been reported previously that most database validation studies assess only three of these parameters, specifically: completeness, accuracy and timeliness [20]. Timeliness in our case was not directly assessed, but we do know that database entry for SQUAD occurs on a continual basis. Upon patient discharge, the SQUAD data entry clerks collect patient charts from the wards; the data are entered into SQUAD; and the charts are then sent to medical records. This suggests that timeliness is not an issue with this specific database. SQUAD contains over 100 variables, some only relevant for certain patients (e.g., patients admitted to the ICU or on the obstetrics service). We only examined 20 of these variables in the surgical population. These 20 variables were agreed upon as the most clinically and administratively important variables. It is likely that the other variables in SQUAD do not have the same degree of validity as the ones highlighted in this paper. Thus, the database is valid only for the 20 variables in question. This validity should not be extrapolated to the variables not specifically addressed in this study.

The future of global surgery hinges upon a solid understanding of the current state of the problem and an accurate way to monitor patients and outcomes over time. SQUAD is an attempt to develop a surgical registry that is appropriate and feasible in the low-resource setting. The current study validates the data captured by SQUAD, rendering it a powerful tool on multiple fronts. We need to develop additional simple, sustainable and valid registries that are easy to roll out across multiple centers in LMICs in order to truly begin to understand, and therefore to improve, the global burden of surgical disease.

Acknowledgements We appreciate the support this project received from the departments of surgery, obstetrics and gynecology and anesthesia at MRRH. We would also like to thank the SQUAD advisory committee as well as the data entry staff for their ongoing dedication. L.I. was supported by the Doris Duke Foundation. Partial funding for this project came from the Massachusetts Medical Society via a grant to GAA.

Compliance with ethical standards

Conflict of interest The authors have no conflict of interest to declare.

References

1. Meara JG, Leather AJM, Hagander L et al (2015) Global Surgery 2030: evidence and solutions for achieving health, welfare, and economic development. Lancet 386:569–624
2. Khuri SF, Daley J, Henderson W et al (1998) The Department of Veterans Affairs' NSQIP: the first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. National VA Surgical Quality Improvement Program. Ann Surg 228:491–507
3. Itani KMF (2009) Fifteen years of the National Surgical Quality Improvement Program in review. Am J Surg 198:S9–S18
4. Ko CY, Hall BL, Hart AJ et al (2015) The American college of surgeons national surgical quality improvement program: achieving better and safer surgery. Jt Comm J Qual Patient Saf Jt Comm Resour 41:199–204
5. Kumar BD, Kumari CM, Sharada MS et al (2012) Evaluation of the medical records system in an upcoming teaching hospital-a project for improvisation. J Med Syst 36:2171–2175
6. Wong R, Bradley EH (2009) Developing patient registration and medical records management system in Ethiopia. Int J Qual Health Care J Int Soc Qual Health Care ISQua 21:253–258
7. Tumusiime G, Was A, Preston MA et al (2016) The quality and utility of surgical and anesthetic data at a Ugandan Regional Referral Hospital. World J Surg. doi:10.1007/s00268-016-3714-8
8. Ttendo SS, Was A, Preston MA et al (2016) Retrospective descriptive study of an intensive care unit at a Ugandan Regional Referral Hospital. World J Surg 40:2847–2856. doi:10.1007/s00268-016-3644-5
9. Meara JG, Greenberg SL (2015) The Lancet Commission on Global Surgery Global surgery 2030: evidence and solutions for achieving health, welfare and economic development. Surgery 157:834–835
10. Quinn JW, Greenberg SLM, Reisel J, Schlesinger L (2015) SQUAD in Uganda: Surgical QUality Assurance Database. The Lancet Commission on Global Surgery Teaching Case
11. Musevani H (2011) State of the Nation Address, Kampala, Uganda
12. Statistics UBo National Population and Housing Census 2014. In: Uganda Ro editor, Kampala, Uganda
13. Mamlin BW, Biondich PG, Wolfe BA et al. (2006) Cooking up an open source EMR for developing countries: OpenMRS—a recipe for successful collaboration. In: AMIA Annual Symposium Proceedings 2006, pp 529–533
14. Bilimoria KY, Liu Y, Paruch JL et al (2013) Development and evaluation of the universal ACS NSQIP surgical risk calculator: a decision aide and informed consent tool for patients and surgeons. J Am Coll Surg 217:833–842
15. Khuri SF, Daley J, Henderson W et al (1997) Risk adjustment of the postoperative mortality rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs surgical risk study 1. J Am Coll Surg 185:315–327
16. Anderson JE, Rose J, Noorbakhsh A et al (2014) An efficient risk adjustment model to predict inpatient adverse events after surgery. World J Surg 38:1954–1960. doi:10.1007/s00268-014-2490-6
17. The World Bank Data Catalog (2016) Washington DC, The World Bank
18. McHugh ML (2012) Interrater reliability: the kappa statistic. Biochem Med 22:276–282
19. Chen H, Hailey D, Wang N et al (2014) A review of data quality assessment methods for public health information systems. Int J Environ Res Public Health 11:5170–5207
20. Chen H, Yu P, Hailey D et al (2014) Methods for assessing the quality of data in public health information systems: a critical review. Stud Health Technol Inform 204:13–18