European Journal of Cardio-Thoracic Surgery 49 (2016) 1470–1475
doi:10.1093/ejcts/ezv385 Advance Access publication 5 December 2015

ORIGINAL ARTICLE

Cite this article as: Salati M, Falcoz P-E, Decaluwe H, Rocco G, Van Raemdonck D, Varela G *et al*. The European thoracic data quality project: An Aggregate Data Quality score to measure the quality of international multi-institutional databases. Eur J Cardiothorac Surg 2016;49:1470–5.

The European thoracic data quality project: An Aggregate Data Quality score to measure the quality of international multi-institutional databases[†]

Michele Salati[a,*], Pierre-Emmanuel Falcoz[b], Herbert Decaluwe[c], Gaetano Rocco[d], Dirk Van Raemdonck[c], Gonzalo Varela[e] and Alessandro Brunelli[f], on behalf of the ESTS Database Committee

[a] Unit of Thoracic Surgery, AOU Ospedali Riuniti, Ancona, Italy
[b] Department of Thoracic Surgery, Nouvel Hopital Civil, Strasbourg, France
[c] Department of Thoracic Surgery, University Hospital, Leuven, Belgium
[d] Department of Thoracic Surgery, National Cancer Institute, Pascale Foundation, Naples, Italy
[e] Department of Thoracic Surgery, University Hospital, Salamanca, Spain
[f] Department of Thoracic Surgery, St James's University Hospital, Leeds, UK

* Corresponding author. via De Gasperi 17/c, 60020 Offagna, Ancona, Italy. Fax: +39-071-5964481; e-mail: michelesalati@hotmail.com (M. Salati).

Received 6 July 2015; received in revised form 15 September 2015; accepted 28 September 2015

Abstract

OBJECTIVES: To describe the methodology for the development of data quality metrics in multi-institutional databases, deriving a cumulative data quality score [Aggregate Data Quality score (ADQ)]. The ESTS database was used to create and apply the metrics. The Units contributing to the ESTS database were ranked for the quality of data uploaded using the ADQ.

METHODS: We analysed data obtained from 96 Units contributing with at least 100 major lung resections (January 2007 to December 2014). The Units were anonymized assigning a casual numeric code. The following metrics were developed for measuring the data quality of each Unit: (i) record Completeness (COM); rate of present variables on 16 expected variables for all the records uploaded [1 – ('null values'/total expected values for the Unit) × 100, the concept of 'null value' was defined for each variable]; (ii) record Reliability (REL); rate of consistent checks on 9 checks tested for all the records uploaded [1 – (valid controls/total possible controls for the Unit) × 100, specific reliability control queries were defined]. These two metrics were rescaled using the mean and standard deviation of the entire dataset and summed, obtaining: (iii) ADQ score: [COM rescaled + REL rescaled]; it measures the cumulative data quality of a given dataset. The ADQ was used to rank the contributors.

RESULTS: The COM of ESTS database contributors varied from 98.6 to 43% and the REL from 100 to 69%. Combining the rescaled metrics, the obtained ADQ ranged between 2.67 (highest data quality) and −7.85 (lowest data quality). Comparing the rating using just the COM value to the one obtained using the ADQ, 93% of Units changed their position. The major change was the drop of 66 positions considering the ADQ list.

CONCLUSIONS: We described a reproducible method for data quality assessment in clinical multi-institutional databases. The ADQ is a unique indicator able to describe data quality and to compare it among centres. It has the potential of objectively guiding projects of data quality management and improvement.

Keywords: Database management systems • Registry • Data quality • Quality indicators

INTRODUCTION

Data collection represents the cornerstone on which to base any knowledge process and the consequent decision-making strategies [1].

The quality of the data collected should be taken into utmost consideration, being able to directly influence the results of analytic models as well as the management of clinical practice [1, 2].

The European Society of Thoracic Surgeons (ESTS) database is a mono-specialistic, multi-institutional and international data collection, aimed at monitoring and improving the quality of care in the Thoracic Surgery specialty across Europe [3].

In 2010, the ESTS database was used to demonstrate the applicability in a medical database of data quality metrics originally developed and derived from other fields (economics and banking) [4].

In the present study, the ESTS Database Committee developed an analytic model to measure the data quality of each Unit contributing to the ESTS database. Two different metrics, data

[†]Presented at the 23rd European Conference on General Thoracic Surgery, Lisbon, Portugal, 31 May–3 June 2015.

completeness and data reliability, were defined and calculated and then aggregated in a unique data quality score, called Aggregate Data Quality score (ADQ). Consequently, it was possible to assign to each contributor an ADQ value and to classify the Units according to the level of the quality of data provided.

This study represents the first attempt to assess the data quality of a multi-institutional, international and mono-specialistic database, such as the ESTS Registry, in our specialty.

This analysis should facilitate future data quality improvement strategies for any single contributor, leading to an increase of the overall data quality of the ESTS database.

METHODS

Partially borrowing concepts and methods already explained and tested in a previous work published in 2010 [4], we defined specific metrics for quantifying the quality of data of a multi-institutional database as the ESTS Registry.

European Society of Thoracic Surgeons registry data analysed

Characteristics of the European Society of Thoracic Surgeons Registry. The ESTS database is a mono-specialistic data collection, gathering information from different sections (i.e. lung surgery, mediastinal surgery, pleural surgery, chest wall surgery) of the thoracic surgery specialty across Europe. The contribution is totally voluntary and each Thoracic Surgery Unit uploads its own data using a common online platform (multi-institutional). All the European Units with at least one staff surgeon as member of the ESTS are allowed to participate to the ESTS database (international). For each procedure uploaded in the database are collected more than 75 variables, almost completely managed as structured data (this means that the numeric or text values are well defined and fixed) with a low grade of variability across years (monolithic database) [4].

Units Inclusion criteria. Data uploaded from those Units contributing to the ESTS database with at least 100 major lung resections from January 2007 to December 2014 were used for the present study. All the Units examined were anonymized and identified using a casual numeric code.

Variables used for the analysis. Among the variables describing each procedure collected (about 75, depending upon the type of procedure), a sample of 16 variables (21%) was selected (Table 1). These variables were chosen taking into account their relevance, being the core variables used for the calculation of the Composite Performance Score (CPS). The CPS is a clinical score developed in general thoracic surgery to measure the quality of care of a Unit, and it represents one of the fundamental parameters to be applied to the ESTS Institutional Accreditation Program [5].

For the present analysis, we used the terms 'record' and 'procedure' as synonyms. In fact, in a typical medical monolithic registry, a record is each single entry represented by a treated patient (or procedure) described by one or multiple values (categorized as variables) [4]. From another perspective, a 'record' is a row of a two-way table reporting data for different 'variables' indicated in each 'column' (Fig. 1).

Data quality metrics developed

The two metrics initially used to evaluate the data quality were the record 'Completeness (COM)' and the record 'Reliability (REL)'.

Table 1: List of variables selected for completeness and definition of the correspondent null values

Variable	Null/missing value
1 BMI	Every blank AND every out-of-limits point. Limits: BMI >5 AND BMI <50
2 AGE	Every blank AND every out-of-limits point. Limits: age = 0 with weight >20, age >0 AND age <110
3 FEV1	Every blank AND every out-of-limits point. Limits: FEV1 >10 AND FEV1 <200
4 DLCO	Every blank AND every out-of-limits point. Limits: DLCO >10 AND DLCO <200
5 ppoFEV1	Every blank AND every out-of-limits point. Limits: ppoFEV1 >10 AND ppoFEV1 <200
6 ppoDLCO	Every blank AND every out-of-limits point. Limits: ppoDLCO >10 AND ppoDLCO <200
7 PREOP INV MED STAG	Filtering the entire dataset for <primary neoplastic malignant>, every blank. Exception: PREOP INV MED STAGING NOT MISSING if morphology = 2
8 CT	Filtering the entire dataset for <primary neoplastic malignant>, every blank. Exception: CT NOT MISSING if morphology = 2
9 PET	Filtering the entire dataset for <primary neoplastic malignant>, every blank. Exception: PET NOT MISSING if morphology = 2
10 CARDIAC COMORBIDITY	Every blank in cardiac comorbidity 1 – (non-blank in cardiac comorbidity 2 after filtering for cardiac comorbidity 1 = blank) – (non-blank in cardiac comorbidity 3 after filtering for cardiac comorbidity 1 and cardiac comorbidity 2 = blank)
11 OTHER COMORBIDITY	Every blank in other comorbidity 1 – (non-blank in other comorbidity 2 after filtering for other comorbidity 1 = blank) – (non-blank in other comorbidity 3 after filtering for other comorbidity 1 and other comorbidity 2 = blank)
12 LUNG EXCISION PROCEDURE	Before filtering the entire dataset already filtered for lung subgroup = <lung excision>, count number of blanks for lung excision procedure
13 MORPHOLOGY	Every blank
14 COMPLICATION	Every blank in complication 1 – (non-blank in complication 2 after filtering for complication 1 = blank) – (non-blank in complication 3 after filtering for complication 1 and complication 2 = blank)
15 OUTCOME AT DISCHARGE	Every blank
16 OUTCOME AT 30 DAYS	Every blank

BMI: body mass index; FEV1: forced expiratory volume in 1 s; DLCO: carbon monoxide lung diffusion; ppoFEV1: predicted postoperative FEV1; ppoDLCO: predicted postoperative DLCO; PREOP INV MED STAG: preoperative invasive mediastinal staging; CT: computed tomography; PET: positron emission tomography.

Combining the COM and REL after a rescaling procedure, it was possible to obtain a third cumulative metric, the ADQ.

Definitions and formulas

Completeness. This parameter measures the extent to which data are not missed and is of sufficient breadth and depth to describe the corresponding set of the real world [6]. In the present analysis, it reflects the rate of present values on the 16 expected variables for all the records uploaded by each Unit.

The formula used to calculate the completeness was: COM = 1 – ('null values'/total expected values for the Unit) × 100. The concept of 'null value' (that represents a missing or an out-of-scale value) was defined for each variable as reported in Table 1.

Reliability. This parameter measures the extent to which data are coherent with each other within the same record (elsewhere this was also indicated as cross-record consistency) [4, 6]. In the present analysis, it reflects the rate of valid checks on the 9 checks defined and tested for all the records uploaded by each Unit.

The formula used to calculate the completeness was: 1 – (valid controls/total possible controls for the Unit) × 100. Specific reliability control queries were defined within a single record, as reported in Table 2.

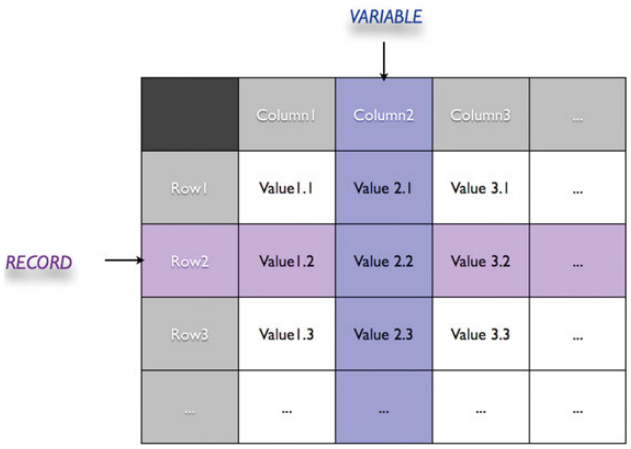

Figure 1: Base of data terminology.

Aggregate Data Quality score. This parameter merges in a single value the two previous metrics. The ADQ measures the global degree of quality for the given amount of data uploaded by each Unit.

It was obtained as: rescaled COM of the Unit + rescaled REL of the Unit.

The COM and REL were rescaled using the following formulas:

Rescaled COM = (COM of the Unit – average COM of all the examined Units)/standard deviation of all the examined Units

Rescaled REL = (REL of the Unit – average REL of all the examined Units)/standard deviation of all the examined Units

Rescaling the original COM and REL allowed summing two parameters that, despite having the same scale (ideally ranging from 0 to 100%), showed a variability of values extremely different. After the correction, the weight of the COM and REL to the cumulative ADQ was perfectly balanced.

Analytic model

For each Unit contributing to the ESTS database, the specific COM was calculated. This parameter was firstly used to rank the Units taking into account just one data quality metric.

Then, the REL was calculated for each Unit and both the COM and REL were rescaled.

In order to verify the potential correlation between the two rescaled metrics, we calculated the Spearman's correlation test.

The rescaled COM and REL were summed to obtain the ADQ metric for each Unit. The ADQ was used to definitely rank the Units according to the cumulative quality of data provided. The two obtained lists were then compared (Fig. 2).

RESULTS

We examined the data uploaded to the ESTS database from 96 European Thoracic Surgery Units that matched the inclusion criteria.

The mean number of records uploaded per Unit was 391, ranging from 101 to 1449. As a whole, the totality of the records

Table 2: List of control queries

Control queries	Definition
1 FEV1 AND ppoFEV1	If FEV1 <ppoFEV1, then inconsistent/no. of patients with FEV1
2 DLCO AND ppoDLCO	If DLCO <ppoDLCO, then inconsistent/no. of patients with DLCO
3 FEV1 AND ppoFEV1	If FEV1 present and ppoFEV1 not present, then inconsistent/no. of FEV1 present
4 DLCO AND ppoDLCO	If DLCO present and ppoDLCO not present, then inconsistent/no. of DLCO present
5 PREOP INV MED STAG AND CT AND PET	If preoperative med stag is blank and CT AND PET are not blank/no. of blank preoperative invasive mediastinal staging
6 PREOP INV MED STAG AND CT AND PET	If preoperative med stag is not blank and CT AND PET are blank/no. of not blank preoperative med staging
7 CARDIAC AND PULMONARY AND CARDIOPULMONARY COMPLICATIONS	If any cardiac or pulmonary complication is 'yes' and cardiopulmonary complications is 'no', then inconsistent/no. of patients with any cardiac or pulmonary complication
8 CARDIAC AND PULMONARY AND CARDIOPULMONARY COMPLICATIONS	If any cardiac or pulmonary complication is 'no' and cardiopulmonary complications is 'yes', then inconsistent/no. of patients without any cardiac or pulmonary complication
9 OUTCOME AT DISCHARGE AND OUTCOME AT 30 DAYS	If outcome at discharge is 'died' and outcome at 30 days is alive/no. of died in hospital

FEV1: forced expiratory volume in 1 s; DLCO: carbon monoxide lung diffusion; ppoFEV1: predicted postoperative FEV1; ppoDLCO: predicted postoperative DLCO; PREOP INV MED STAG: preoperative invasive mediastinal staging; CT: computed tomography; PET: positron emission tomography.

uploaded by the examined Units counted for about the 90% of all the records (major lung resection procedures) collected within the ESTS database in the same period.

Table 3 presents part of the COM and REL values obtained for each Unit (the integral version of the list in the Supplementary Material). The Units were ranked using the descending COM value. The COM value ranged from 98.6% for the Unit 00-51 to 43% for the Unit 07-55. The mean COM for the entire examined dataset was 79.5% and the mean REL was 98.8%.

Table 4 reports part of the rescaled COM and REL values and the obtained ADQ for each Unit (the integral version of the list in Supplementary Material).

We did not find an association between the rescaled COM and REL (Spearman's correlation coefficient: −0.17, $P = 0.10$).

The ADQ ranged between 2.67 for the Unit 00-51 that reflected the highest value of cumulative data quality for a Unit and −7.85 for the Unit 13-03 (that had the lowest one).

For each Unit, the definite position in the list sorted using the ADQ was reported and shown side by side to the one previously assigned. The last column of the table describes the direction of the movement within the list. Comparing the rating using the COM to the one obtained using the ADQ, 93% of Units changed their position. The Units that have moved in the lists showed a change of eight positions on average.

The global distribution of the ADQ level for all the ranked Units is shown in Fig. 3. From the 36th Unit are registered negative values of ADQ.

DISCUSSION

Context

The system of knowledge that ultimately leads every decision-making process is based on an initial collection of data [7]. In the current big data era [8], even the most sophisticated analytic model will not be able to deliver valuable informations if the mass of analysed data is characterized by an intrinsic low data quality level [9]. Therefore, it would be appropriate to estimate the quality of data that we are going to examine, before proceeding in a study. Unfortunately, the vast majority of published scientific papers do not report the quality of the analysed dataset, even if it only means measuring the completeness of the most relevant variables.

In 2011, the ESTS Database committee published a first study focused on the subject of data quality assessment in medical registries [4]. The authors developed potential metrics for measuring the data quality, borrowing them from other fields (such as economics or banking). The obtained data quality indicators (completeness, correctness and consistency) were then applied to the ESTS database, showing a prototype of data quality assessment.

The present study represented a further exploration in the data quality subject a standardized and reproducible analytic model

Figure 2: Flowchart of the analytic model. ADQ: Aggregate Data Quality score.

Table 3: Completeness and reliability value for each Unit

Unit code	COM %	REL %
00-51	98.59	99.60
01-50	94.94	99.46
00-52	94.18	99.52
04-40	95.66	98.74
03-61	93.59	98.57
00-53	90.26	99.58
04-41	90.27	99.51
01-51	89.69	98.85
01-52	90.40	98.84
00-54	86.84	99.42
04-42	87.00	99.22
04-43	87.83	98.23
08-80	84.43	99.28
02-61	82.65	99.64
03-62	82.54	99.36
04-44	81.36	99.75
01-53	80.54	99.63
07-70	81.70	98.74
05-80	77.44	99.82
03-63	80.55	98.70
06-70	89.82	95.50
01-54	78.40	99.29
01-55	80.57	98.53
11-01	79.70	98.21
04-45	81.06	97.56
01-56	74.36	99.05
10-01	88.27	94.28
09-20	75.93	97.47
12-70	71.92	98.49
11-02	69.33	99.35
07-71	76.89	96.64
13-51	63.88	99.72
23-61	62.81	99.80
33-70	64.69	99.09
50-50	62.10	99.62
50-51	60.38	100.00
13-50	62.75	99.10
33-71	60.61	99.60
13-49	60.93	99.49
23-60	64.31	98.16
... stopped at 40 of the 96 Units tested		
Mean	79.5	98.8
SD	11.2	1.2

Mean and SD refer to all the Units examined.
COM: completeness; REL: reliability.

Table 4: Comparison of metrics and ranking of Units

Unit code	COM %	REL %	Rescaled COM	Rescaled REL	ADQ	Rank by ADQ	Rank by COM	Move	Number of moved positions
00-51	98.59	99.60	2.34	0.33	2.67	*1*	*1*	=	0
01-50	94.94	99.46	2.07	0.30	2.37	*2*	*3*	↑	1
00-52	94.18	99.52	2.02	0.32	2.33	*3*	*4*	↑	1
04-40	95.66	98.74	2.12	0.15	2.27	*4*	*2*	↓	2
03-61	93.59	98.57	1.97	0.11	2.08	*5*	*5*	=	0
00-53	90.26	99.58	1.73	0.33	2.06	*6*	*8*	↑	2
04-41	90.27	99.51	1.73	0.31	2.04	*7*	*7*	=	0
01-51	89.69	98.85	1.69	0.17	1.86	*8*	*10*	↑	2
01-52	90.40	98.84	1.74	0.09	1.83	*9*	*6*	↓	3
00-54	86.84	99.42	1.48	0.29	1.77	*10*	*14*	↑	4
04-42	87.00	99.22	1.49	0.25	1.74	*11*	*13*	↑	2
04-43	87.83	98.23	1.55	0.04	1.59	*12*	*12*	=	0
08-80	84.43	99.28	1.30	0.26	1.57	*13*	*15*	↑	2
02-61	82.65	99.64	1.17	0.34	1.51	*14*	*16*	↑	2
03-62	82.54	99.36	1.16	0.28	1.45	*15*	*17*	↑	2
04-44	81.36	99.75	1.08	0.36	1.44	*16*	*19*	↑	3
01-53	80.54	99.63	1.02	0.34	1.36	*17*	*23*	↑	6
07-70	81.70	98.74	1.10	0.15	1.25	*18*	*18*	=	0
05-80	77.44	99.82	0.79	0.38	1.17	*19*	*27*	↑	8
03-63	80.55	98.70	1.02	0.14	1.16	*20*	*22*	↑	2
06-70	89.82	95.50	1.70	−0.55	1.15	*21*	*9*	↓	12
01-54	78.40	99.29	0.86	0.27	1.13	*22*	*26*	↑	4
01-55	80.57	98.53	1.02	0.10	1.12	*23*	*21*	↓	2
11-01	79.70	98.21	0.96	0.03	0.99	*24*	*24*	=	0
04-45	81.06	97.56	1.06	−0.11	0.95	*25*	*20*	↓	5
01-56	74.36	99.05	0.57	0.21	0.78	*26*	*31*	↑	5
10-01	88.27	94.28	1.58	−0.81	0.77	*27*	*11*	↓	16
09-20	75.93	97.47	0.68	−0.13	0.56	*28*	*30*	↑	2
12-70	71.92	98.49	0.39	0.09	0.48	*29*	*32*	↑	3
11-02	69.33	99.35	0.20	0.28	0.48	*30*	*33*	↑	3
07-71	76.89	96.64	0.75	−0.31	0.45	*31*	*28*	↓	3
13-51	63.88	99.72	−0.20	0.36	0.16	*32*	*37*	↑	5
23-61	62.81	99.80	−0.27	0.38	0.10	*33*	*38*	↑	5
33-70	64.69	99.09	−0.14	0.22	0.09	*34*	*35*	↑	1
50-50	62.10	99.62	−0.33	0.34	0.01	*35*	*40*	↑	5
50-51	60.38	100.00	−0.45	0.42	−0.03	*36*	*47*	↑	11
13-50	62.75	99.10	−0.28	0.23	−0.05	*37*	*39*	↑	2
33-71	60.61	99.60	−0.44	0.33	−0.10	*38*	*45*	↑	7
13-49	60.93	99.49	−0.41	0.31	−0.10	*39*	*44*	↑	5
23-60	64.31	98.16	−0.17	0.02	−0.14	*40*	*36*	↓	4
... stopped at 40 of the 96 Units tested									

The italic font is used just to put emphasis into the tho lists.
COM: completeness; REL: reliability; ADQ: aggregate data quality score.

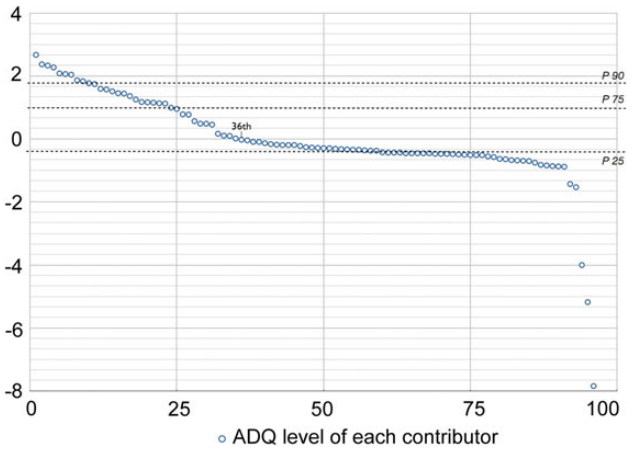

Figure 3: ADQ distribution of the ranked Units. Caption: P 90: 90th percentile, P 75: 75th percentile, P 25: 25th percentile. ADQ: Aggregate Data Quality score.

for measuring the quality of data uploaded by each Unit contributing to the ESTS database.

Main finding

We developed a couple of data quality indicators represented as completeness and reliability, which can be merged to obtain a third metric, the ADQ, able to express in a single value the quality level of a given dataset. This could be considered a model for assessing data quality especially in those registries collecting data from multiple contributors, such as the ESTS database. The methods used to derive the metrics were strictly defined in order to create an analytic model for data quality assessment that is stable and reproducible over the years. It was possible to calculate the rate of completeness and reliability for all the 96 Units enrolled in the study, obtaining some interesting information:

(i) The completeness showed a higher level of variability (max Com: 98.6%, min COM: 43%, COM SD: 11.2) in comparison to the level of reliability (max REL: 100%, min REL: 69%, REL SD: 1.2).
(ii) The overall level of reliability of the ESTS database for the selected data of the enrolled Units was sensibly higher than the level of completeness (mean REL for the entire dataset: 98.8% vs mean COM for the entire dataset: 79.5%).
(iii) The two metrics, even after the rescaling procedure, were not related, contributing independently to the data quality of the ESTS database.
(iv) The ranking process of the Units using firstly a single parameter of data quality, such as the completeness value, and then using a multi-parametric data quality score such as the ADQ, allowed one to obtain two different lists. The comparison of the two classifications showed that:
(v) The ADQ is a more exhaustive indicator and it causes a movement in the list (up or down) for more than 90% of the units;
(vi) The Units below the 35th Unit presented negative values of ADQ. This means that more than 60% of the Units had completeness or reliability or both metric values below the mean for the entire population.

Clinical inference

International registries are instruments of knowledge extremely useful in offering a large amount of data for performing analyses and deriving information. Nowadays, the evidences based on large registries heavily influence the clinical practice and management strategies [10, 11]. Nevertheless, the quality of data collected within these databases is rarely investigated, being complex to measure and difficult to improve, as it is influenced by multiple contributors.

The present study offers an example of an analytic model to measure data quality in large, international and multi-institutional registries, such as the ESTS database. We developed two specific data quality metrics (completeness and reliability), able to measure different aspects of the amount of data uploaded by each ESTS database contributor. Then, we combined these parameters to obtain the ADQ. This indicator offers the possibility of defining in a single value the quality of data of each contributor to the ESTS database. As a consequence, we can:

(i) assess the level of ADQ of a Unit and compare it to the rest of contributors;
(ii) monitor the level of ADQ of a Unit over the years;
(iii) implement actions to improve the data quality of a Unit and verify the results with further ADQ measurements.

Moreover, the presented model could be used for assessing the level of data quality before performing analysis based on multi-institutional data collection. In fact, the overall rate of completeness and reliability could be measured, and those contributors with an ADQ below a predefined level could be excluded from the analysis.

Limitations

(i) As it occurs in most large and multi-institutional registries, the process of data quality assessment was performed in the base of data after the uploading procedure. In other words, the data quality verification process was not intended to validate or improve the original data source.
(ii) At present, the ESTS cannot assess whether the data uploaded by each contributor are equivalent to the ones recorded in the original source (accuracy data quality metric) [4]. This kind of quality control is usually performed only in institutional registries or, at least, in *ad hoc* databases created by a few contributors. Nevertheless, in the context of the accreditation programme, the ESTS is also engaged in activities of accuracy verification for those units considered eligible for accreditation.
(iii) The ADQ obtained for each Unit resulted from the completeness and reliability assessment of the selected group of 16 variables, considered essential for the purpose of the ESTS database because they are used for the calculation of the CPS. Changes in the ADQ value of a Unit could be noted in case of a different process of variable selection.

Recommendation

The use of analytic models for the assessment of data quality should be included in the processes of database management, as well as before proceeding into the analytic phase of any scientific study.

The ADQ and the single data quality indicators (completeness and reliability) presented here should be verified on a regular basis in order to monitor the level of data quality in large databases, especially in international and multi-institutional registries.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *EJCTS* online.

Conflict of interest: none declared.

REFERENCES

[1] Salati M. Reasons to participate in European Society of Thoracic Surgeons database. J Thorac Dis 2015;7(S2):S112–7.
[2] Ferraris VA, Edwards FH, Shahian DM, Ferraris SP. Risk stratification and comorbidity. Card Surg Adult 2008;3:199–246.
[3] Falcoz PE, Brunelli A. The European General Thoracic Surgery Database project. J Thorac Dis 2014;6(S2):S272–5.
[4] Salati M, Brunelli A, Dahan M, Rocco G, Van Raemdonck DEM, Varela G *et al*. Task-independent metrics to assess the data quality of medical registries using the European Society of Thoracic Surgeons (ESTS) Database. Eur J Cardiothorac Surg 2011;40:91–8.
[5] Brunelli A, Berrisford RG, Rocco G, Varela G; on behalf of the European Association of Thoracic Surgeons Database Committee. The European Thoracic Database project: composite performance score to measure quality of care after major lung resection. Eur J Cardiothorac Surg 2009;35:769–74.
[6] Batini C, Cappiello C, Francalanci C, Maurino A. Methodologies for data quality assessment and improvement. ACM Comput Surv 2009;41:1–52.
[7] Fayyad U, Piatetsky-Shapiro G, Smyth P. The KDD process for extracting useful knowledge from volumes of data. Commun ACM 1996;39:27–34.
[8] Calvert M, Thwaites R, Kyte D, Devlin N. Putting patient-reported outcomes on the 'Big Data Road Map'. J R Soc Med 2015; doi:10.1177/0141076815579896.
[9] Wang RY, Strong DM. Beyond accuracy: what data quality means to data consumers. J Manage Inform Syst 1996;12:5–33.
[10] Burt BM, Kosinski AS, Shrager JB, Onaitis MW, Weigell T. Thoracoscopic lobectomy is associated with acceptable morbidity and mortality in patients with predicted postoperative forced expiratory volume in 1 second or diffusing capacity for carbon monoxide less than 40% of normal. J Thorac Cardiovasc Surg 2014;148:19–29.
[11] Falcoz PE, Puyraveau M, Thomas PA, Decaluwe H, Hürtgen M, Petersen RH *et al*. Video-assisted thoracoscopic surgery versus open lobectomy for primary non-small-cell lung cancer: a propensity-matched analysis of outcome from the European Society of Thoracic Surgeon database. Eur J Cardiothorac Surg 2016;49:602–9.