

# Maintenance of a European database

Stefano Passani<sup>1,2,3</sup>, Danilo Pellicano<sup>3,4,5</sup>, Enrico Passariello<sup>3</sup>, Marco Peticarà<sup>3,6</sup>

<sup>1</sup>La Sapienza Medical School, Rome, Italy; <sup>2</sup>City University London, London, UK; <sup>3</sup>KData Clinical srl, Rome, Italy; <sup>4</sup>University School of Computer Engineering Federico II, Naples, Italy; <sup>5</sup>Alma Mater Studiorum University, Bologna Italy; <sup>6</sup>Computer Technology La Sapienza, Rome, Italy  
Correspondence to: Stefano Passani, MD, MHM. KData Clinical srl, Rome, Italy. Email: s.passani@kdataclinical.it.

Submitted Dec 01, 2017. Accepted for publication Apr 16, 2018.

doi: 10.21037/jtd.2018.04.101

View this article at: <http://dx.doi.org/10.21037/jtd.2018.04.101>

## Introduction

Although neither Oxford or Cambridge Dictionaries offer a definition for database (DB), DB maintenance in clinical informatics jargon means the organisation of dedicated operations to keep a clinical research DB “fit for purpose” i.e., serving effectively the research and teaching objectives is designed for, over the number of years that is expected to run, in full compliance with data protection legislation.

A standard DB set-up project should include protecting the data integrity, securing the correct enrolment of those who have the right to be, backing up data, setting up and up-dating disaster recovery processes, facilitating a periodic dataset review with leading clinicians, implementing their advice on the live DB, and all of it in compliance with current European data protection law. The collaborative agreement to support the European Society of Thoracic Surgeons (ESTS) Database has additional features: we also have introduced additional functions including secure third party data imports and data extraction, and introduced a DB Quality Monitor [called Clinical Care Analysis (CCA)] which informs each user about their real time percentage of data completeness and composite performance score (CPS); they are both mandatory items to qualify for the ESTS accreditation process (1).

In this contest we will examine and discuss the set of operations we routinely perform on behalf of the ESTS for their European database, a multi centre, multi-national, procedure-specific multi registries, prospective, population-based longitudinal DB.

And may be in some distant future there will be a standard dictionary item for DB maintenance.

## Background

### ESTS

ESTS DB was founded in 2001 by the ESTS Database committee with the aim to develop risk-adjusted instruments for assessing the performance of thoracic surgery units across Europe. Since then ESTS has made use of its central DB to support the publication of the first risk-adjusted multinational risk-score for mortality (2); based on this initial data ESTS also published a critical comparison of the performance of different units (3). The second version of the DB was launched online in July 2007 with a more extensive Dataset (list of questions and validated answers) and has so far accrued a total of over 120,000 procedures from 205 general thoracic surgery Centres from Europe; also from 16 Brazilian, 1 Canadian, 2 Chinese Contributing Centres (4,5).

### European Union Data Protection Regulation (EUDPR)

Since 24<sup>th</sup> October 1995 there has been a Data Protection Directive 95/46/EC that regulates the processing of personal data. Technology and the way we use data has evolved much faster than legislation, and in 2012 the European Union (EU) Commission started a radical review and up-date of EUDPR, focusing on a more prescriptive legislative proposal; this was approved in Dec. 2015, adopted in Jan. 2016 and is due to be fully enforceable on 24<sup>th</sup> May 2018, after 2 years of post-adoption grace.

The key changes introduced by this more stringent Data Protection Legislation are summarized in *Table 1*.

**Table 1** Summary of European Union Data Protection Regulation (EUDPR) up-grade

---

|   |
|---|
| Greater protection for EU citizens, as it has a much wider applicability; it applies to all companies processing personal data of subjects residing in the Union, regardless of the company's location. It states key responsibilities and accountabilities with the figures of data controller, data processor and data protection officer |
| Penalties, under GDPR organizations in breach of it can be fined up to 4% of annual global turnover or €20 million (whichever is greater)   |
| Consent, consent must be clear, and provided in a clear and plain language. It must be as easy to withdraw consent as it is to give it  |
| Data subject rights, it spells clearly some rights (breach notification, right to access, right to be forgotten, portability, and privacy by design of IT systems)  |

---

EU, European Union; GDPR, General Data Protection Regulation.

## KData Clinical

KData Clinical experience in setting up and running multi-centre international DBs starts in early 1990's with some current KData people involved in health management research work based at Bupa Head Office, London, a British health insurer and hospital provider.

We had access to over 25 years health insurance claims about healthcare provided in about 400 international public and private hospitals both in UK and abroad to an insured population of over 3 million UK-overseas members; the team published original health management research papers about "propensity to treat" analysis, the cost of learning curve in new procedures (laparoscopic cholecystectomy), and relationship between outcomes and costs (6,7). This unique large data denominator allowed also the assembly of a first attempt to profile clinicians' activities by comparing some key indicators of each with the distribution norm of the same available parameters in a chosen disease area (8-10).

The limitations of these activities were essentially that all our analyses were from health Insurance claims data, and therefore contained limited clinical details. Since 2009 we were appointed to take over the running of the ESTS Database from another provider. Since then the DB has been very successful, as shown in *Figure 1*.

It has also supported the production of numerous original research papers, and the running of the calculations of the CPS (11); recently its data has contributed to the international data pool to support up-dating of the TNM system for thymic tumors (12-14).

The ESTS Database is based on a large body of diverse experiences and its methodologies have been successfully tested outside Europe as it hosts not only European, but also Brazilian, Canadian and Chinese data; DB maintenance relies on a timely and meaningful liaison between researchers, ESTS Database director & team and DB keepers (KData Clinical); therefore effective communication

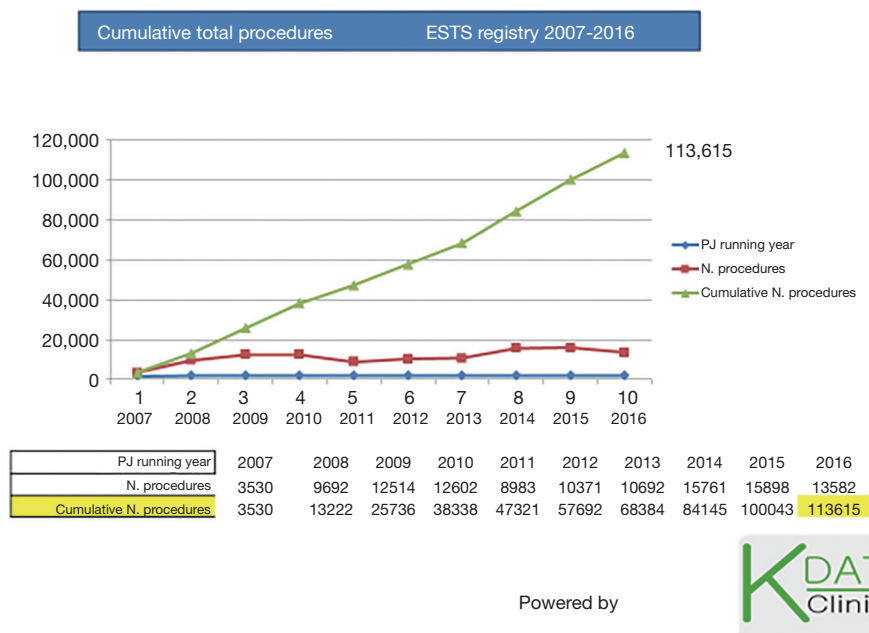
and team work are key ingredients to its success.

## **"Nuts and Bolts" of running a International, multi-centre, multi-procedure prospective, longitudinal, population-based clinical outcomes registry**

In this context we will examine the "Nuts and Bolts" of running successfully an international DB (ESTS), which is made of multiple registries (5 in total).

Currently the ESTS Database aims to fulfil the following objectives:

- (I) To facilitate suitable clinical data collection enabling to set current best clinical practice standards about the core (lung) and satellite clinical areas (thymoma, mesothelioma, NETs, chest wall).
- (II) To support the ESTS accreditation process.
- (III) To supply suitable data for specific projects of clinical research.
- (IV) Obviously a wide multi-expertise team is involved in the many steps to achieve these objectives. As KData we include the following responsibilities and tasks as "database maintenance".
- (V) Secure hosting of DB and access to it.
- (VI) Design and implementation of technical platform to run the DB on, and its up-keeping by periodic up-grade releases, all in compliance with EU data protection legislation.
- (VII) Safe keeping of data.
- (VIII) Secure accessing to the DB either to collect data into it and to retrieve it out of it.
- (IX) Processes in compliance with current European data protection legislation.
- (X) Close liaison with the DB director about near



**Figure 1** Growth of ESTS Database from 2007 to 2016. ESTS, European Society of Thoracic Surgeons.

future directions, problem solving; datasets definitions and clinical content revisions.

- (XI) Enrolment of new contributors and creation and maintenance of their personal accounts.
- (XII) Implement up-dating clinical contents of each registry upon request of the DB director and his/her clinical leads.
- (XIII) Design and implementation of real-time data analysis; called CCA, it is pertinent to status and clinical area of each user.
- (XIV) Data retrieval and file extraction on behalf of authorized users and DB director.
- (XV) Data import from third party DBs.
- (XVI) Preliminary data analysis for the yearly ESTS DB report.
- (XVII) Review and finalising ESTS report content under the guide of the DB director.

### Key elements of DB maintenance

#### Collaborative approach

A collaborative integration of clinical research and its system requirements is recognized as a better approach to designing, building, supporting a large, multi-centre, international DB, particularly for prospective, population-

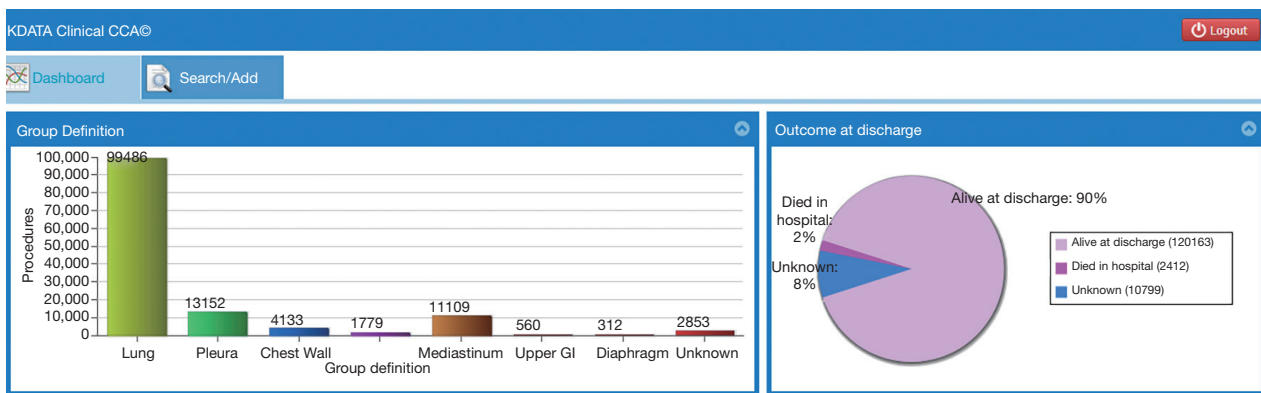
based, longitudinal, observation studies; interaction between clinicians and clinically-minded informatics introduces a certain energy or “tension” (I don’t believe it, I don’t understand it) which needs to be converted into operational consensus to result in greater clarity and purpose, and ultimately to its success. Mastering effective communication between two “tribes” with such diverse mindsets can be quite challenging, but is of paramount importance, and relies on individuals who can understand, reconcile and manage effectively the interface between the two tribes.

#### Datasets

As datasets we mean a registry list of headings, questions and replies.

Clinically comprehensive datasets about chosen clinical areas and/or tracker procedures needs to be agreed by the ESTS Database Committee, under the director guidance. An ideal dataset has to be designed:

- (I) To provide data useful to with stratify patients according to severity of illness using agreeable algorithms;
- (II) To profile changes in clinical practice over time;
- (III) To explore aspects of professional training;
- (IV) To examine both mortality and morbidity as short and long term outcomes.



**Figure 2** CCA or ESTS Database data summary. CCA, Clinical Care Analysis; ESTS, European Society of Thoracic Surgeons.

Increasingly other end points are considered:

- (I) Percentage of data completeness;
- (II) Cost relation to effectiveness & quality;
- (III) Appropriateness of treatment.

Currently ESTS has 5 datasets in use:

- (I) Lung (the core registry);
- (II) Thymoma (satellite);
- (III) Mesothelioma (satellite);
- (IV) Neuro endocrine tumors (NETs) (satellite);
- (V) Chest wall (satellite).

The process of launching a new dataset requires to propose an “idea” to the DB director; if approved after consultation with the DB committee, the proponent can act as the clinical leader of that particular registry and has the responsibility to secure some initial funding and support from her/his peers. ESTS usually supports the yearly cost of up-keeping a new registry. Once a preliminary dataset has been assembled, usually KData gets involved to support its finalisation and integration into the “live” DB. Each data item will have a definition and its validation criteria.

### Contributing process

It prescribes the minimum requirements to collaborate to the ESTS Database.

Each contributor is to be an “active” ESTS member (yearly ESTS membership’s quote paid for) and should collect all variables of the agreed dataset for each and every patient. The process to enlist new contributors requires to complete a form from the ESTS homepage (<http://www.ests.org>), that will be vetted by the ESTS Database director and executive director; once approved KData receives a request to proceed to create a new account, and release own

personal login and password details to the new contributor.

Data should be collected prospectively on the ESTS Database; to improve data completeness there is also a “minimum dataset” the complete fulfilment of which still secures that procedure has being recorded into the DB, and it is valid for CPS purposes. However ESTS recommends that the whole dataset is fulfilled as much as possible, as it would be much more useful for clinical research.

Yearly third-party DB retrospective data imports are available, once there has been a suitable matching of donor and recipient datasets to secure a meaningful data transfer. ESTS has funded most data imports in the past, but it is unlikely that will continue in view of increasing number of them.

Merging data centrally is KData’s responsibility and the “custodians” are ESTS and KData.

KData doesn’t own any of the data that it keeps and has not right whatsoever over ESTS Data.

Data is anonymously reported, independently accessed and encrypted to other users.

Participation to the DB project is totally free and voluntary, and strongly recommended by ESTS. In fact participation to the ESTS Database with at least 150 major lung resections for at least 3 consecutive years is one of the key requirements for the ESTS accreditation program.

Access the DB from ESTS website or by using the url address: <https://ests.kdataclinical.it>.

Contributor will be able to visualize their up-to-date summary of surgical activity; this feed-back process is called CCA, and includes a few surgical activity indicators [total n procedures, types lung procedures, video-assisted thoracoscopic surgery (VATS), outcome at discharge, CPS and eligibility for ESTS accreditation], as shown in *Figure 2*.

### Resources

ESTS DB director, ESTS executive director, KData DB & analysis system, ESTS DB team, KData DB team.

### People

A guess-estimate rather than actual man/days per year to manage successfully the ESTS Database data runs in the many hundreds; thankfully most of these valuable “resources” are ESTS members who support the process by voluntary efforts, resulting in significant on-going costs containment. However financial resources are an essential asset to set up and run well a DB. For the KData part, our team is made of one web developer, one data analyst, an account manager and a clinician (ex., cardio-thoracic surgeon), who all interact with the ESTS Database team on periodical and as-required basis.

### System

A suitable web based data collection system, natively (created for, at source) as fully compliant with data protection laws, is the minimum information system requirement: KData has developed its own proprietary platform, now in its third generation; an integral part of it is feed-back system integrated in the main DB system (called CCA); it is updated in real time with each data added to the DB; so your latest data is represented in a fixed set of data analysis that include relevant parameters such total n of procedures, age, type of surgical procedure, morbidity and mortality outcomes, percentage of data completeness; other specific items are added to profile better each of the five registries. It is registered “compliant” with current EU data protection regulation.

### Data export

Data collected in a system will need to be exported to any contributor: the current routines allows each recognized ESTS contributor to export out of the system his own data, without the need for external intervention. He/she can also apply to the ESTS DB director to have an anonymised copy of the whole ESTS data denominator for research purposes: the current ESTS process regulating it needs a proposal containing the research topic; this is presented and discussed at ESTS council, and in event of being approved, data is released to the named researcher.

### Data import

Data import from third party DBs is also offered by the

ESTS to those contributors who run their own local systems. KData supports this process by a standard operating procedure (SOP).

### Data quality

KData gets involved in the data validation process since the beginning of a new ESTS DB design: the early integration of clinical and informatic validation standards results in a smoother implementation of a new registry either as a stand-alone or as integrated into the existing DB; it also helps with the DB overall data quality by reducing bias. At “working party level” each data item in the proposed dataset is reviewed and given a clear data definition and data type in line with established conventions. Once agreed and signed-off, KData can design and implement a suitable clinical registry system.

Once up and running the best start for achieving good quality data is to try to secure good levels of data completeness and accuracy. Both parameters are statutory requirements the ESTS needs for a centre to become Eligible for the ESTS accreditation process.

This process must be open to external scrutiny, as it happens for financial accounts. Contributing to the ESTS Database is an essential requirement and during an on-site visit part of the ESTS accreditation process, data from local clinical notes is checked against data entered into the ESTS Database. Key check points are data completeness, accuracy and consecutivity.

### Reporting: CCA

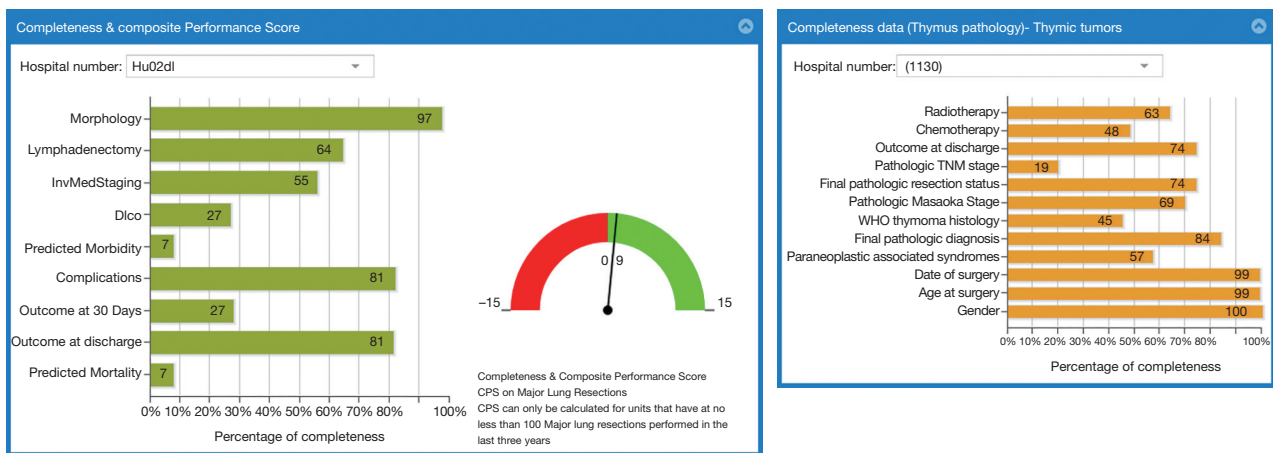
An immediate feed-back mechanism, it up-dates its self as data is added to the system. Currently offers a real-time percentage of data completeness per registry, as not all contributors use all five ESTS registries; also auto-calculates the CPS per contributing unit, and guides the clinical team on which areas of their registry need improving, as shown in *Figure 3*.

### Conclusions

An international DB running smoothly generates a significant number of assets for clinical research, teaching, quality assurance and healthcare improvements. It must be “compliant” with current and near future data protection legislation.

DB maintenance appears to be an essential set of tools





**Figure 3** CCA: ESTS composite performance score & percentage completeness for thymic tumors registry. CCA, Clinical Care Analysis; ESTS, European Society of Thoracic Surgeons.

that underpins the correct, uneventful and productive functioning of the any DB, including the ESTS one.

### Acknowledgements

None.

### Footnote

*Conflicts of Interest:* S Passani is Managing Director of KData. The other authors have no conflicts of interest to declare.

### References

1. Brunelli A, Berrisford RG, Rocco G, et al. 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.
2. Berrisford R, Brunelli A, Rocco G, et al. The European Thoracic Surgery Database project: modelling the risk of in-hospital death following lung resection. *Eur J Cardiothorac Surg* 2005;28:306-11.
3. Brunelli A, Varela G, Van Schil P, et al. Multicentric analysis of performance after major lung resections by using the European Society Objective Score (ESOS). *Eur J Cardiothorac Surg* 2008;33:284-8.
4. ESTS Database Committee. Database Annual Report—2017 NB: Report available to ESTS Members only. Available online: <http://www.ests.org/>
5. SBCT Database Committee. First Database Annual Report—2017 NB: Report available to SBCT Members only.
6. Passani S, Garrad-Cole P, Bailey A, et al. 4th International meeting of the Society of Minimal Invasive Orlando, USA. 10/94, “Learning Curve of Laparoscopic Cholecystectomy: Effects on Total Procedure Cost”. Available online: <http://www.ismit.org/conferences/former-smit-conferences.html>
7. Passani S, Garrad-Cole P, Bailey A, et al. 10th Annual meeting of the International Society of Technology Assessment in Health Care, Baltimore, Maryland, USA. 1994, “The Effects of Introducing a Laparoscopic Cholecystectomy in a Stable Population of British Privately Insured Patients with regards to Consumption of Health Resources for Gall Bladder Disease”. Available online: <https://uia.org/s/or/en/1100014760>
8. Passani S. Go with the Wind: A Systems Approach to Health Care Planning. In: McClean S, Millard P. editors. *Clinical Care Analysis: A Tool for Clinicians Profiling*. 1996: Chapter 1, part 2:35-41. ISBN 1-85315-277-3.
9. Passani S. Quantitative Modelling in the Management of Health Care Conference 1994, University of Manchester, UK. “Clinical Care Analysis: A Tool for Clinicians Profiling”.
10. Passani S. Society of Cardiothoracic Surgeons of Great Britain and Ireland Annual Conference 1995, Loughborough University. “Clinicians Profiling: Can the NHS Learn from the Private Sector”. Available online: <https://scts.org/annual-meeting/programme-archives/>
11. Brunelli A, Falcoz PE, D’Amico T, et al. European

- Guidelines of Structure and Qualifications of General Thoracic Surgery. *Eur J Cardiothorac Surg* 2014;45:779-86.
12. Kondo K, Van Schil P, Detterbeck FC, et al. The IASLC/ITMIG Thymic Epithelial Tumors Staging Project: proposals for the N and M components for the forthcoming (8th) edition of the TNM classification of malignant tumors. *J Thorac Oncol* 2014;9:S81-7.
  13. Nicholson AG, Detterbeck FC, Marino M, et al. The IASLC/ITMIG Thymic Epithelial Tumors Staging Project: proposals for the T Component for the forthcoming (8th) edition of the TNM classification of malignant tumors. *J Thorac Oncol* 2014;9:S73-80.
  14. Detterbeck FC, Stratton K, Giroux D, et al. The IASLC/ITMIG Thymic Epithelial Tumors Staging Project: proposal for an evidence-based stage classification system for the forthcoming (8th) edition of the TNM classification of malignant tumors. *J Thorac Oncol* 2014;9:S65-72.

**Cite this article as:** Passani S, Pellicano D, Passariello E, Perticarà M. Maintenance of a European database. *J Thorac Dis* 2018;10(Suppl 29):S3542-S3548. doi: 10.21037/jtd.2018.04.101