Databases in tracheal diseases

Michele Salati[1], Benoit Jacques Bibas[2]

[1]Unit of Thoracic Surgery, AOU Ospedali Riuniti Ancona, Ancona, Italy; [2]Division of Thoracic Surgery, Heart Institute (InCor), Hospital das Clinicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil

Correspondence to: Michele Salati. AOU Ospedali Riuniti Ancona, Via Conca 71, 60126, Ancona, Italy. Email: michelesalati@hotmail.com.

Abstract: During the last two decades surgery has considerably changed not only under the pressure of technological advances but also for the wide dissemination of a clinical practice based on the data driven approach. In fact, surgeons became increasingly used to analyze and to review their own result in order to optimize the subsequent clinical management with the aim of offering a better quality of care to their patients. Obviously, this virtuous process must be founded on reliable and comprehensive collections of clinical data. The detailed storage of information about patients, treatments and results is even more important for those categories of disease that are less common and where the standardization of practice is poor. Tracheal diseases submitted to surgical treatment are indeed rare and the description of different aspect of clinical practice is fragmented within the scientific literature. For this reason, the needs exist of a standardized and shared data repository with a multi-institutional and international connotation that could represent a benchmark for collecting data about tracheal diseases surgery and at the same time a valuable instrument for improving the quality of care in this field.

Keywords: Database; surgery; tracheal disease

International databases for tracheal diseases treatment: status of the art

In a period when the surgical practice and the patient management is more and more driven by the data science (1), it is very difficult to find examples of public data repositories aimed at collecting surgical procedures for treating tracheal diseases.

In particular, if we consider the topic of tracheal stenosis, that is one of the most investigated subject in the field of tracheal disease from several point of view (epidemiology, etiology, risk factors, surgical treatment, non-surgical management, results, follow up, cost, quality of life, etc.), we find that in the last two decades just 1,600 papers have been published worldwide (2). Among them, the analysis for extracting knowledge and science has never been based on databases with public access, super-institutional and specifically conceived for collecting information about surgical procedures for treating tracheal diseases.

As a matter of fact, the two largest international thoracic surgery databases, as the Society of Thoracic Surgeons general thoracic surgery database (https://www.sts.org/registries-research-center/sts-national-database/general-thoracic-surgery-database) and the European Society of Thoracic Surgeons database (http://www.ests.org/collaboration/default.aspx), even though currently offer to the contributors the chance of uploading their data about tracheal diseases and related procedures, they don't provide a specific section for an exhaustive collection of information about this topic, especially in terms of outcomes and follow up.

Therefore, a greater effort should be made in order to create a common data repository, that could ideally be public, on-line and international, for gathering information about tracheal disease and treatments, able to improve the knowledge and the quality of care offered to our patients affected by this condition.

The need of a multi-institutional data collection for tracheal diseases surgical procedures

The implementation of a new database always requires

efforts at different levels: from the theoretical planning to the funding acquisition to cover costs, passing through the definition of data flows, the contributors identification and the achievement of technical solution for the interface development and for the data storage (3). Considering in particular a public registry, the complexity of this process is multiplied by the multi-institutional and potentially international nature of the database.

On the other side of the scale, the realization of a common data repository offers several benefits and advantages in comparison to single institution-maintained databases. This is especially important for a surgical database about tracheal diseases, where the clinical activity of a single hospital is usually characterizes by a low workload which results in a low number of procedures collected per time.

Some specific benefits provided by an international data collection for tracheal diseases surgical procedures could be:

(I) Standardization of terminology. According to the socio-cognitive theory of terminology, a term is a natural language representation of a unit of understanding, considered relevant to given purposes, applications, or groups of users (4). Unfortunately, due to the fact that the scientific papers published in peer reviewed journals about tracheal surgery are mainly based on mono-institutional databases, the terminology used to collect tracheal surgical procedures data is often not consistent among different papers. This negatively influence the chance of comparing the activities of two or more units in terms of epidemiological characteristics of the treated patients, surgical procedures performed and achieved outcomes. Moreover, both for qualitative and quantitative data, even though the information for a specific subject are collected using the same terms, the descriptive categories (qualitative data) as well as the unit of measurements (quantitative data) could be different. A shared multi-institutional database will force the contributors to speak exactly the same language. The adoption of a standardized common terminology for describing procedures will ensure the comparison among units, will facilitate the creation of multi-institutional studies and will optimize the knowledge extraction inherent in tracheal surgery field (5).

(II) Indirect standardization of clinical practice. A surgical database is a powerful instrument able to capture in a rigorous and systematic way punctual information about a specific medical piece of world. This information will allow to better understand several aspects of that world leading our clinical and surgical actions with a higher level of effectiveness and quality (6,7). Obviously, the process of data influence on surgery, that is more and more described in the theories of data driven practice, needs a constant, prospective and accurate collection of every single descriptor (that we would call "variable" in a database) of every single phase (preoperative evaluation, operative treatment, postoperative course and follow up) that makes up the entire clinical pathway of a patient. As a consequence, the participation of a data collection process requires an effort of clinical practice standardization in order to build a solid and reliable data repository, where data will be stored with high completeness and accuracy levels, ensuring the quality of the following steps of analysis and interpretation of data.

(III) The power of large numbers. Looking at the scientific literature about tracheal surgery is glaringly obvious that the published works describe for the most part the experience of a single center, retrospective and based on small cohort study (8). This reflects the nature of the clinical characteristics of tracheal diseases that are managed through surgical procedure quite infrequently and in centers where the tracheal surgery procedures represent a minority in the general thoracic surgery practice. The consequence of this scenario is that the analysis and the extraction of information and knowledge about tracheal surgery are based on a paucity of data. This problem is exponentially greater if we consider the analysis of those events that are rare in this kind of surgery as the complications or the negative functional results and the mortality following the surgical treatment. For these reasons, the creation of a multi-institutional database with a deep description of the tracheal surgery practice founded on punctual and shared variables will increase the ability of studying several aspects of this subject. In fact, the collection of data from a multitude of contributors, once standardized and verified in terms of quality, allow to overcome the limits inherent in the analyses of events that occur with low frequency.

(IV) Setting benchmarks. Taking into consideration the three previous point and emphasizing the super-institutional connotation of an international tracheal surgery database, we can easily understand that it could play a pivotal role for increasing the quality of care of patients affected by tracheal diseases. In fact, speaking the same language to describe the tracheal surgery in terms of patients' characteristics, operative procedures and consequent results as well as sharing a more standardized clinical practice among different centers, it seems extremely facilitated the different experience comparison among several centers. The analysis of epidemiology, risk factors, surgical indications, surgical techniques, early and late results and patients' follow up can be analyzed as a whole or compared at different levels, selecting specific cohorts of patients. This will result in a clearer definition of activity benchmarks, setting desirable standards of care in the tracheal surgery scenario.

Characteristics of an international database for tracheal diseases

An electronic database is a virtual repository that collects information about a specific subject of the world. The characteristics of the repository as well as of the stored data can be very different leading to instruments able to describe this subject with a huge spectrum of depth and accuracy.

For instance, we can build a database where the information is clearly collected in a totally structured way or, at the opposite side, as text with undefined format. In both cases, the database will be able to offer information and knowledge about that subject, but all the phases of imputation, cleaning and transformation of data will require not comparable efforts (9).

In order to have a database as effective as possible:

(I) When we build a multi-institutional database the process of definition and selection of the variables should be very rigorous. In particular considering the tracheal disease we do not have many other preexisting public registries that can be used as point of reference. At the same time, considering the scientific literature addressing the subject of tracheal surgery, the papers are often not aligned and standardized in the adoption of variables for describing characteristics of the patients,

procedures and outcomes.

(II) The database should have a very high proportion of structured variables. In fact, once the subject and the purpose of a data collection is clearly defined, the use of semi-structured or unstructured data can be minimized. This will allow to optimize the process of knowledge extraction from the database because the procedures of data cleaning and data analysis are extremely simplified by the use of structured data. Such an effort is particularly beneficial when we consider rare pathological conditions and treatments, as in the case of tracheal surgery, because a structured collection of data prevents possible loss of information.

(III) Especially considering databases that collect data from multiple institution and multiple nations, it seems very useful the identification of a subset of variables within the entire dataset which are mandatory. These are a group of variables strictly linked to the final purpose of that specific data collection and essential for the description of the patients, for the qualification of the performed procedure and for the observed outcomes. These are usually labeled as "core variables". Missing one of these variables will be equivalent to miss the information of the related entire procedure, that, as a consequence, shouldn't be uploaded within the database itself. The definition of the core variables could increase the effective informative power of the collected data as well as the data quality and the procedure of data quality management.

(IV) The use of an online platform offers the chance of accessing the electronic database from any place at any time. This represents a prerequisite to facilitate a prospective data upload. It is demonstrated that a prospective data collection increases the quality of information gathered within the database and positively influence its timeliness (see below).

(V) The strongest difference between a database for tracheal disease and the other international registries developed for collecting data about general thoracic surgery is that these ones are built taking into consideration mainly diseases with an oncological perspective. As a consequence, the preoperative characteristics of the patients as well as the surgical treatment and the results are weighted on the basis of oncological criteria and survival. Differently, a tracheal registry should be

Transl Cancer Res 2020;9(3):2149-2153 | http://dx.doi.org/10.21037/tcr.2020.02.32

2152

Salati and Bibas. Collecting clinical data in tracheal surgery

more focused on the restoration of physiological functions as breathing, speaking and swallowing. An innovative vision should permeate the collection of preoperative, intraoperative and postoperative data with the aim of obtaining information able to generate units of measurements of functional outcomes.

(VI) Particular attention should be paid to the data collection within the follow-up section. This is due to the fact that tracheal surgery patients usually undergo several evaluations over time after the treatment. Moreover, the functional outcomes may have different degrees of evaluation and not just a binary option as for the mortality in oncologic surgery. This implies that the data collected at any follow-up time must be clearly organized with metrics of outcome that trace the same ones registered in the first postoperative functional evaluation.

(VII) A multi centric database should finally optimize the timeliness, that is a data metric which measure the ability of a database of returning information about the real world that it should describe. This is obviously affected by how fast the information system state is updated after the real-world system change. Considering this point of view the database should adopt the simplest technical solutions for gathering data, should offer to the end-users a clear interface for uploading the information about a new procedure and possibly stimulate the end-user himself to a very updated data entry. This last point is complex to obtain for multi-institutional databases and could be encouraged by offering to the contributors prompt feedback and analysis about their own data (in form of graph or dashboards) and related quality.

Future perspective

Taking into account the peculiarity of tracheal diseases as well as the evolving scenario of tracheal surgery characterized by high complexity and rapid technical advances, the need of a multi-institutional database for tracheal operative procedures is obvious.

During the last two years the European Society of Thoracic Surgeons (ESTS) and the Brazilian Society of Thoracic Surgery (SBCT) started a joint project in order to create a registry for collecting information about airways

surgery, as reported during the last meetings of the two Societies.

In particular, a committee with SBCT and ESTS representatives built a shared program for:

- ❖ Defining the database architecture;
- ❖ Selecting the variables of interest (preoperative characteristic of the patients, airways disease descriptors, surgical procedure characteristics, clinical and functional outcomes, follow up);
- ❖ Defining the variables in a dedicated dictionary;
- ❖ Defining the data flows;
- ❖ Defining the end user interface and the rules for data upload;
- ❖ Securing the funding for the project.

At present, the developed beta version of this registry is going through the planned one-year testing phase, where several Units worldwide are uploading cases in order to verify its content, structure and different technical aspects. Within the next year the on-line final version of the database should be officially launched.

The final goal of this international project is the creation of a dedicated airways surgery registry that could contribute to extract information, standardize practice and increase the quality of care for patients affected by tracheal diseases and submitted to operative treatments.

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editors (Benoit Jacques Bibas, Paulo Francisco Guerreiro Cardoso and Konrad Hoetzenecker) for the series "Recent Developments in Benign Tracheal Stenosis" published in Translational Cancer Research. The article was sent for external peer review organized by the Guest Editors and the editorial office.

Conflicts of Interest: The series "Recent Developments in Benign Tracheal Stenosis" was commissioned by the editorial office without any funding or sponsorship. BJB served as the unpaid Guest Editor of the series. The other authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related

to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Maier-Hein L, Vedula SS, Speidel S, et al. Surgical data science for next-generation interventions. Nat Biomed Eng 2017;1:691-6.
2. Farzanegan R, Feizabadi M, Ghorbani F, et al. An overview of tracheal stenosis research trends and hot topics. Arch Iran Med 2017;20:598-607.
3. Gupta P, Mata-Toledo RA, Monger MD. Database Development Life Cycle. JISOM 2011;5:8-17.
4. Temmerman R. Why traditional terminology theory impedes a realistic description of categories and terms in the life sciences. Terminology 1998;5:77-92.
5. Fernandez FG, Falcoz PE, Kozower BD, et al. The Society of Thoracic Surgeons and the European Society of Thoracic Surgeons general thoracic surgery databases: joint standardization of variable definitions and terminology. Ann Thorac Surg 2015;99:368-76.
6. Wade DT. Assessment, measurement and data collection tools. Clin Rehabil 2004;18:233-7.
7. Fayyad U, Piatetsky-Shapiro G, Smyth P. The KDD process of extracting useful knowledge from volumes of data. Commun ACM 1996;39:27-34.
8. Siciliani A, Rendina EA, Ibrahim M. State of the art in tracheal surgery: a brief literature review. Multidiscip Respir Med 2018;13:34.
9. Batini C, Cappiello C, Francalanci C, et al. Methodologies for data quality assessment and improvement. ACM Comput Surv 2009;41:1-52.