

Large databases: additional value or inherent risk?

Journal of Hand Surgery
(European Volume)
2018, Vol. 43(4) 453-454

© The Author(s) 2018



Reprints and permissions:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/1753193417754182

journals.sagepub.com/home/jhs



Large databases are collections and storage systems for a huge amount of population-based or regional system-wide data. The large databases in medicine contain personal information, insurance details, symptoms, surgical procedures, complications of treatments, implants used, and many others. Busy surgeons sometimes make use of the large databases for research. A review in orthopaedic research showed a significant increase in database use over 10 years from January 2006 to December 2015 (Weinreb et al., 2017).

There are two major categories of databases: *administrative* and *clinical*. The *administrative* databases are created with the help of billing information and are not primarily used to answer research questions. The *clinical* databases consist of patient records and clinical information, which enables the researcher to investigate clinical questions. These databases are either national (public) or private. An example of a national administrative database is the National Inpatient Sample (NIS) in the United States. The private analytics databases derive the data from their clients, such as hospitals or insurance companies, who pay for the use. Examples for clinical databases include the American College of Surgeons National Surgical Quality Improvement Programme (ACS NSQIP) and the Norwegian Arthroplasty Register (NAR). Usually a fee is required to obtain the data from some of the national or private databases.

Strengths and possible pitfalls in using the databases

Due to the high numbers of patients included in databases, their data can be considered representative of populations that are of interest. Their data can reliably give information on procedure volumes or length of hospital stay. With the help of administrative data, health care utilization or outcome differences due to geographical differences or patient demographics can be assessed.

Limitations are their financial and administrative restrictions, which lead to differences in detail and accuracy of the data sets, often caused by differences

in the quality of coding or underreporting of events. Often, clustering in the data set of a specific diagnosis can lead to incorrect data. The use of multivariate regression models is advisable in order to check for clustering and thereby avoid misleading conclusions. Another common problem is that very small differences between groups can result in statistically significant differences due to the large sample number in these databases, but these differences may not be clinically meaningful. Therefore, clinical relevance and absolute, as well as relative, differences between groups have to be taken into account in order to judge findings adequately. The researchers must define a meaningful clinically significant difference threshold prior to the analysis of the data.

Between different administrative databases from the same geographical region, there may be differences if they are used to answer the same study questions. Bohl et al. (2014) found considerable differences in conclusion on lumbar spine fusion procedures in two large American databases. The study results can be affected by which database an investigator uses.

The clinical databases are relatively smaller registries that mostly collect data prospectively for a particular diagnosis or procedure, such as procedure-specific outcomes (complications, functional scores, patient-satisfaction or X-rays, etc.). These databases may define complications or comorbidities differently, which leads to heterogenous data. Madigan et al. (2013) found that study results from these databases can vary from statistical significance in either direction depending on the database used.

Points of care and advice in using the databases

Inaccuracy or bias in conclusions in reports based on large databases have been described in other specialties (Bohl et al., 2014; van Walraven and Austin, 2012), such as bias in patient selection and over-interpretation of significance of authors' findings. We noted these problems in the work submitted to the Journal as well. Several recent submissions that

I reviewed presented interesting results at first sight. A closer look later brought problems to the surface, which included the following.

1. *Inadequate raw data*: Inability to provide most or all the raw data for running statistical tests by the reviewers, which renders the whole statistical work-up questionable. We could not check authors' statistical methodologies. Therefore, we were unable to confirm whether statistical significance in the report is proper.
2. *Missing inclusion criteria*: For example, if the percentage of a clinical manifestation in patients with carpal tunnel syndrome in a region is analysed, it is necessary to know if the database includes all patients with the diagnosis of carpal tunnel syndrome and not only those who have been treated operatively. If the patients being managed conservatively or having refused surgery are missed, the results will be wrong.
3. *Missing exclusion criteria*: For instance, if the database only includes patients with a certain health care insurance and excludes those with another insurance or no insurance at all, it is not possible to test the prevalence of carpal tunnel syndrome in a population. This mistake will result in a too high prevalence of carpal tunnel syndrome in the statistical calculation.

I believe the authors should at first make sure to use the database properly, then correct statistics should be used to aid analysis. Finally significance of the differences found in the study should be carefully defined and interpreted. Conclusions should be made according to clinically meaningful differences rather than purely statistical differences.

On reading a report based on large databases, the readers should look critically at inclusion and exclusion criteria in the reports, statistical methods used, and whether the reported results address important clinical questions properly and ultimately add to evidence-based knowledge. Authors should also be aware of the fact that the conclusions obtained based on large databases, that is the findings in a region or a country, may be very different from the readers' practice; sometimes the conclusions are not useful for those readers' daily practice.

References

- Bohl DD, Russo GS, Basques BA et al. Variations in data collection methods between national databases affect study results: a comparison of the nationwide inpatient sample and national surgical quality improvement programme databases for lumbar spine fusion procedures. *J Bone Joint Surg Am.* 2014, 96: e193.
- Madigan D, Ryan PB, Schuemie M et al. Evaluating the impact of database heterogeneity on observational study results. *Am J Epidemiol.* 2013, 178: 645–51.
- van Walraven C, Austin P. Administrative database research has unique characteristics that can risk biased results. *J Clin Epidemiol.* 2012, 65: 126–31.
- Weinreb JH, Yoshida R, Cote MP, O'Sullivan MB, Mazzocca AD. A review of databases used in orthopaedic surgery research and an analysis of database use in arthroscopy: *The Journal of Arthroscopic and Related Surgery.* Arthroscopy. 2017, 33: 225–31.

Mike Ruettermann

Department of Plastic Surgery, University Medical Centre Groningen, University of Groningen, Groningen, The Netherlands and HPC – Oldenburg, Institute of Hand and Plastic Surgery, Oldenburg, Germany
Email: mikeruettermann@yahoo.com