Heart, Lung and Circulation (2016) **25,** 237–242
1443-9506/04/$36.00
http://dx.doi.org/10.1016/j.hlc.2015.01.006

REVIEW

Application of Clinical Databases to Contemporary Cardiac Surgery Practice: Where are We now?

Akshat Saxena, MBBS, MS [a,b*], Andrew E. Newcomb, MBBS [c], Vikrant Dhurandhar, MBBS [a,b,d], Paul G. Bannon, MBBS, PhD [a,b,d]

[a]Department of Cardiothoracic Surgery, Royal Prince Alfred Hospital, Sydney, NSW, Australia
[b]Discipline of Surgery, Sydney Medical School, The University of Sydney, Sydney, NSW, Australia
[c]Department of Cardiothoracic Surgery, St Vincent's Hospital Melbourne, Melbourne, Vic., Australia
[d]The Baird Institute, Sydney, NSW, Australia

Received 6 January 2015; received in revised form 10 January 2015; accepted 13 January 2015; online published-ahead-of-print 19 February 2015

Cardiac surgery has embraced and encouraged the use of large, multi-institutional datasets in clinical practice. From a research perspective, database studies have facilitated an increased understanding of cardiac surgery. Among other uses, they have allowed an investigation of disease incidence and mortality, high risk groups, disparities in health care delivery and the impact of new devices and techniques. Databases are also important tools for clinical governance and quality improvement. Despite their obvious utility, clinical databases have limitations; they are subject to treatment bias, contain missing data and cannot establish causality. Moreover, the ongoing maintenance of the database requires significant human and financial resources. In the future, inclusion of more detailed follow-up data and integration with other datasets will improve the utility of clinical databases.

Keywords Cardiac surgery • Database • Society of Thoracic Surgery (STS) • Society for Cardiothoracic Surgery (SCTS) • Australia and New Zealand Society of Cardiac and Thoracic Surgeons (ANZSCTS) • Observational studies

Background

Whilst prospective and well-conducted randomised controlled trials (RCTs) are regarded as the highest level of research evidence in clinical medicine they often cannot be performed due to ethical, financial or practical reasons [1,2]. This is particularly true in surgery where only 20-30% of primary treatment interventions are supported by randomised evidence [3]. For thoracic surgery only 14% of treatments are supported by randomised evidence [4]. Consequently, the effects of new devices and techniques are often based on observational data obtained from large databases. Database studies can also be used to study rare diseases, interventions, adverse events and side-effects and to examine whether results from RCTs translate into effective real world practice. Support from randomised trials cannot and should not be demanded of all treatment interventions used in clinical practice [5]. Performing high-quality observational studies requires the availability of large datasets with clinically important variables. Cardiac surgery has benefitted from the widespread use of large, multi-institutional datasets in clinical practice. Beyond their value for research, these datasets are also used for quality control and clinical governance [6,7]. Moreover, these datasets have been used to generate risk assessment tools which have significant utility for patient counselling, operative decision-making and treatment allocation [8–11]. Overall, the widespread adoption of international datasets in cardiac surgery has facilitated the improved outcomes that have been observed in the past two decades [12,13]. Herein are discussed various aspects of large databases relevant to cardiac surgery clinical research and practice.

*Corresponding author at: Department of Cardiothoracic Surgery, Royal Prince Alfred Hospital, Missenden Road, Camperdown, New South Wales, Australia, 2050. Tel.: +61 433890393, Email: akshat16187@gmail.com

Advantages of Clinical Databases

The research utility of large databases is indisputable. One particular advantage is the ready availability of data on a large patient sample. This generally represents a larger proportion of the actual patient population and subsequently reduces sampling error and improves external validity [14]. Large databases also capture data on patients with rare diseases or those undergoing an infrequently performed procedure. Often, in clinical medicine, RCTs cannot be performed because low disease incidence or uncommon adverse outcomes would require extremely large sample sizes. In these situations, large observational database series may allow us to infer causal relationships particularly if good study designs and statistical modelling (e.g. propensity matching techniques) are applied [5,15,16]. Furthermore, RCTs are expensive; one study showed an average cost of $USD 12 million [17]. Also, because RCTs typically take several years before being published the data they contain is withheld from the medical field and may be less timely and relevant at the time of publication. In light of these disadvantages, a good database study, rather than a RCT, may represent a more appropriate research tool in certain clinical circumstances [5].

Databases have other advantages. The inclusion of a large number of patients from multiple centres reduces procedure selection biases and improves validity over single centre studies. Furthermore, because databases capture a heterogeneous sample of the population, they reflect real world practice (clinical effectiveness) rather than the effect of interventions in ideal circumstances (clinical efficacy) [14]. Moreover, because databases utilise data that has already been collected, studies based on databases are usually less expensive, less obtrusive, quicker to perform and less questionable ethically. Databases also allow for ongoing review of disease incidence and mortality, trends in the utilisation of different interventions and disparities in health care. In cardiac surgery, the ongoing assessment of patient data has demonstrated to clinicians and healthcare policy-makers the shifting profile of patients undergoing cardiac surgery. They have shown that surgical patients are older and have more co-morbidities [12]. Concurrently, analysis of data from databases has been critical in demonstrating the improved outcomes in several high-risk groups, such as the elderly [18,19]. This has subsequently facilitated increased acceptance of surgical intervention in these patients which has translated into superior long-term survival. The fact that trainee surgeons can safely perform cardiac surgery was demonstrated through database analysis [20–22]. Moreover, the impact of surgical volume on outcomes, disparities in outcomes on the basis of socioeconomic status and comparisons in outcomes on the basis of operative techniques have all been facilitated by database analyses. Many of the findings reported in database studies have subsequently been investigated through large, multi-randomised trials (Tables 1 and 2).

Well-conducted cardiac surgery databases have also facilitated the development of clinically useful risk-assessment tools. Risk tools such as EuroScore, Society of Thoracic Surgeons Score and AusScore have all been developed using well-designed, prospective datasets [8–11]. These tools are useful for several reasons. Firstly, they inform patients and families about their surgical risk. In clinical practice, this risk is estimated from the clinicians' knowledge and experience, but risk-prediction models can provide an objective and individualised probability estimate of an adverse outcome, such as operative mortality. Hence, risk models, derived from databases can facilitate informed consent in patients [6,23].

Secondly, risk models aid in quality control. In-hospital or 30-day mortality are often used as a surrogate markers for quality of care. The use of these markers alone, however, is confounded by the fact that there is considerable variation in disease severity, co-morbidity and demographic profiles between individual patients and across institutions [24]. Consequently, comparison of operative outcomes without adjusting for prognostically influential variables is invalid. Databases, through collection of prognostic variables, allow more valid analysis to be performed. Using risk models at an institutional or population level and comparing the estimated risk with actual outcomes is useful for quality assurance and identifying underperforming centres. Thirdly, risk scores assist with surgical and interventional decision-making. For example, risk models are now used to determine whether patients are better suited for transcutaneous aortic

Table 1 Advantages of Clinical Databases.

Advantages of Clinical Databases	
1	Research utility -Large sample size -Study rare disease or uncommon adverse outcomes -Study infrequently performed procedures -Studies less expensive, obtrusive than RCTs -Studies less ethically questionably than RCTs -Studies are quicker to perform
2	Facilitate development of risk-assessment tools
3	Drive quality improvement and clinical governance
4	Ongoing review of disease incidence, disease mortality, volume-outcome relationships, national trends in the use of procedures and disparities in health care

Table 2 Disadvantages of Clinical Databases.

Disadvantages of Risk Scoring Systems	
1	Subject to treatment bias
2	Sampling error and missing data
3	Significant human and financial resources required
4	Regular reviews and updates to ensure database remains contemporaneous