SHORT REPORT

Accuracy of Clinical Data in a Population Based Vascular Registry

J. Laustsen,[1]* L. P. Jensen,[2] A. K. Hansen[3] and
The Danish National Vascular Registry

Departments of [1]Cardiothoracic and Vascular Surgery T, Aarhus University Hospital, Skejby Sygehus, Aarhus, [2]Vascular Surgery B, Gentofte University Hospital, Gentofte, and [3]Vascular Surgery V, Aarhus University Hospital, Aalborg Sygehus, Aarhus, Denmark

Introduction. *Clinical databases are increasingly being employed to evaluate the quality of treatments, including patients with peripheral vascular disease. Valid data is vital to the value of these analyses.*
Objective. *To assess the validity of clinical data in a population-based national vascular registry.*
Design. *Traditional reproducibility study was supplemented by refilling of data by an independent observer, thereby creating three data sets for comparison.*
Materials and methods. *Twenty prospectively recorded electronic forms from each department were selected randomly from the Danish National Vascular Registry. Data forms were refilled by the surgeons of the department concerned, and by an independent member of the board of the Danish National Vascular Registry. Refilling was performed blinded to the original forms.*
Conclusions. *A high degree of accuracy of clinical data can be achieved. An independent observer makes it possible to evaluate the classification of observer dependent parameters and explain differences in the reproducibility of data.*

Key Words: Vascular surgical procedures; Registries; Databases; Human; Medical audit; Medical records.

Introduction

The usefulness of clinical databases depends on the quality of the data included. Combining data incorporation from various electronic sources, and the possibility of automated continuous check and filtration of errors will ensure completeness and validity of some of the data.[1]

On the other hand much data are obtained by some kind of clinical assessment which will be less valid. Indication for treatment (diagnosis), procedure performed, risk factors and complications are examples of the latter.

It was the aim of this study, to describe the validity of these observer dependent data in a population-based[2] national vascular registry.

Material and Methods

Data on 20 operations from each of the 10 Danish vascular surgery departments where randomly drawn from the Danish National Vascular Registry, 10 operations for abdominal aortic aneurysms and 10 infra inguinal operations for occlusive disease. The departments were asked to re-complete the data sheets, and this procedure was then repeated by an independent member of the board of the Danish National Vascular Registry. The refilling of forms was done on the basis of patient's notes and all other available information, but blinded to the primary data set.

Three theoretically identical data sets were created for analyses. Data were matched and transferred to SPSS for analysis.

Cross tabulation analyses for risk factors and complications were performed at 2 levels;

1. Exact reproducibility.

---

*Corresponding author. Jesper Laustsen, Department of Cardiothoracic and Vascular Surgery T, Aarhus University Hospital, Skejby Sygehus, DK-8200 Aarhus N, Denmark.

2. Binary reproducibility (present/absent).

The classification of indication for surgery, risk factors and outcome measures is based on a simple multiple-choice system.[3] Coding of operations was performed according to the NOMESCO Classification of Surgical Procedures (NCSP).[4]

Results

The reproducibility of the coding of the operation performed was 90% on a six digit level. If analyzed on a five digit level the accuracy was 100%. The accuracy of the coding for indications for surgery (diagnosis) was improved considerably when the classification was simplified i.e. CLI or claudication and for AAA rupture or not. This improved the reproducibility from 76 to 97%.

There was no difference in reproducibility between the three data sets concerning coding for operations and indication for surgery (diagnosis). Accuracy of redo vs. board, i.e. two vascular surgeons interpreting exactly the same data, was not superior.

Reproducibility of coding for risk factors and complications on the other hand varies considerably between the departments. As demonstrated in Fig. 1

the reproducibility of risk factors and complications is enhanced by simplifying the different data points into a more robust clinical classification i.e. whether the risk factor or complication is present or not (binary data).

By adding the third data set from an independent observer the poorer reproducibility in, for example, department 4, has to be due to bad primary coding whereas it in department 5, it is poor re-coding performed by the actual department that explains the results (Fig. 2).

An interesting observation is that no 'massage of data' could be recognized, as there was no tendency to aggravate the indications or over code the risk factors.

Discussion

The lack of a gold standard limits the validation of vascular registries to reproducibility studies against other databases, existing paper records or refilled data forms.[5–7] A study of an orthopaedic database suggested that some data such as complications may be more difficult to acquire or enter than diagnosis or type of operation, perhaps because complications imply an admission of failure.[8] Yeoh studied the
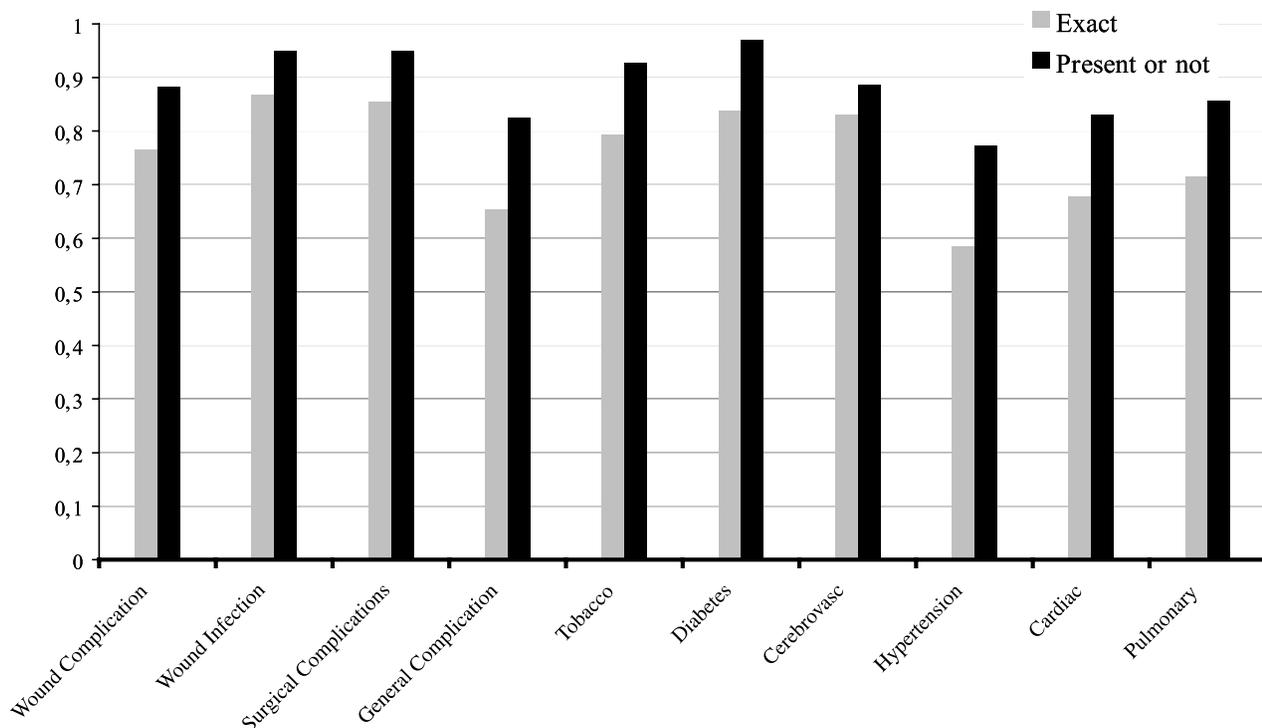

Fig. 1. Reproducibility of coding for complications and risk factors for exact match, and reproducibility when the parameter is classified as present or not.
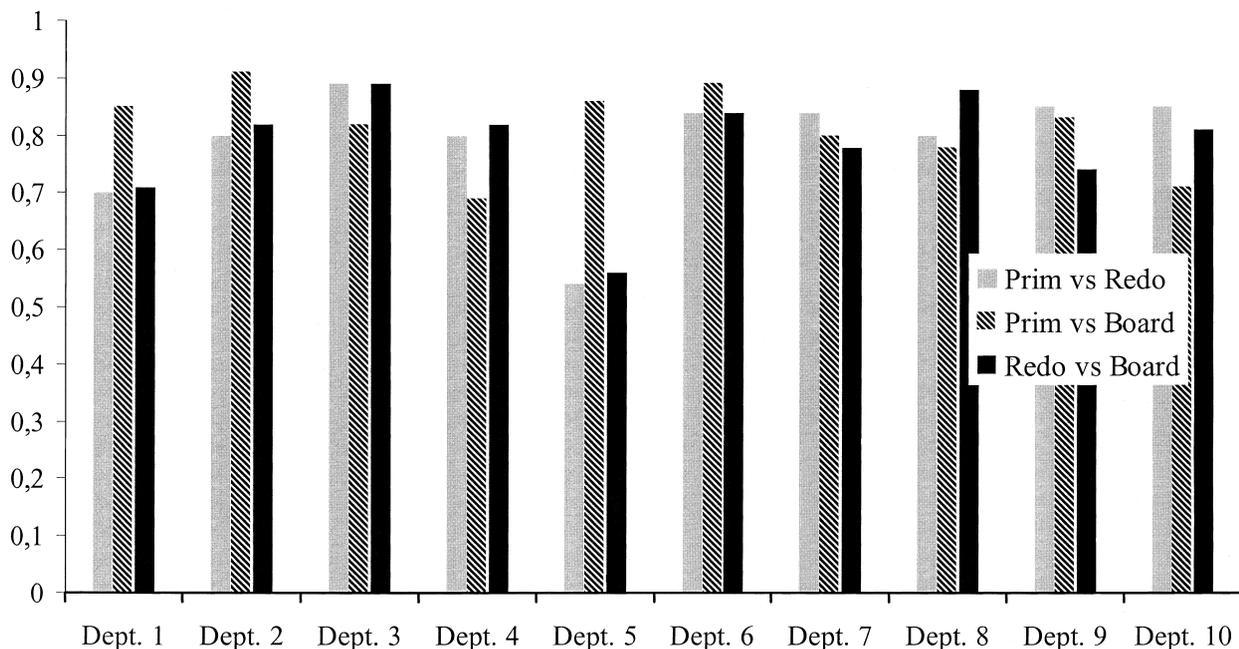
Fig. 2. Variation of reproducibility of coding for complications among the 10 centres performing vascular surgery in Denmark. The three data sets compared pair-wise for exact match. Prim is the primary data set from the vascular registry. Redo is the data set re-abstracted by the departments. Board is the data set created by the independent board member.

accuracy and completeness of pediatric diagnoses. When the coding was performed by doctors the accuracy was 85%.[9] Other studies strengthen that the achievable accuracy when data is generated by clinicians approaches 90%.[10-12] Fine and colleagues found that the database of the Society for Cardiothoracic Surgeons of Great Britain and Ireland was incomplete and unreliable on many variables. After a period of monitoring, validation, and feedback in 10 centres, completeness of the database improved significantly, but the overall reliability of the data improved only marginally.[13] Our results confirm that softer data such as risk factors are less accurate than information on operative procedures or diagnoses. But if the data points are made robust a satisfactory accuracy can be achieved in a large-scale registry. The validity of data is influenced by local conditions including staff structure. It is, therefore, mandatory to document completeness and reproducibility of data before comparisons of outcome is performed.

Lack of improvement in reproducibility between the data sets from refilled forms and the data sets obtained from an independent observer classifying exact the same information must depend on weakness in the classification of the parameters in question.

We supplemented traditional reproducibility studies with an independent observer and we

were able to evaluate the classification of our parameters and explain differences in the reproducibility of data.

References

1 POWSNER SM, WYATT JC, WRIGHT P. Opportunities for and challenges of computerisation. *Lancet* 1998; **352**: 1617–1622.
2 OLSEN J, BASSO O, SORENSEN HT. What is a population-based registry? *Scand J Public Health* 1999; **27**: 78.
3 JENSEN LP, SCHROEDER TV, MADSEN PV, LORENTZEN JE. Vascular registers in Denmark based on personal computers. *Ann Chir Gynaecol* 1992; **81**: 253–256.
4 The Nordic Medico-Statistical Committee. NOMESCO Classification of Surgical Procedures. Uppsala and Copenhagen, 2000.
5 BERGQVIST D, TROENG T, ELFSTROM J, HEDBERG B, LJUNGSTROM KG, NORGREN L et al. Auditing surgical outcome: ten years with the Swedish Vascular Registry—Swedvasc. The Steering Committee of Swedvasc. *Eur J Surg Suppl* 1998; 3–8.
6 KANTONEN I, LEPANTALO M, SALENIUS JP, FORSSTROM E, HAKKARAINEN T, HUUSARI H et al. Auditing a nationwide vascular registry—the 4-year Finnvasc experience, Finnvasc Study Group. *Eur J Vasc Endovasc Surg* 1997; **14**: 468–474.
7 LEPANTALO M, SALENIUS JP, LUTHER M, YLONEN K. Introduction of a population-based vascular registry: validity of data and limitations of registration. The Finnvasc Study Group. *Br J Surg* 1994; **81**: 979–981.
8 BARRIE JL, MARSH DR. Quality of data in the Manchester orthopaedic database. *BMJ* 1992; **304**: 159–162.
9 YEOH C, DAVIES H. Clinical coding: completeness and accuracy when doctors take it on. *BMJ* 1993; **306**: 972.
10 WYATT JC. Clinical data systems, part 3: development and evaluation. *Lancet* 1994; **344**: 1682–1688.
11 JICK H, JICK SS, DERBY LE. Validation of information recorded on

general practitioner based computerised data resource in the United Kingdom. *BMJ* 1991; **302**: 766–768.

12 Nazareth I, King M, Haines A, Rangel L, Myers S. Accuracy of diagnosis of psychosis on general practice computer system. *BMJ* 1993; **307**: 32–34.

13 Fine LG, Keogh BE, Cretin S, Orlando M, Gould MM. How to evaluate and improve the quality and credibility of an outcomes database: validation and feedback study on the UK Cardiac Surgery Experience. *BMJ* 2003; **326**: 25–28.