

Existing Data Resources as Reference Standards for a Clinical Data Warehouse

Jonathan S. Einbinder, MD, MPH¹, Kenneth W. Scully, MS¹, Robert G. Sawyer, MD²

¹Department of Health Evaluation Sciences and ²Department of Surgery

University of Virginia Health System, Charlottesville, VA

BACKGROUND

Reliable, objective gold standards for the evaluation of informatics applications are often not available. As a result, review of patient records by human experts may be used as a reference standard.¹ However, record review can be expensive and time consuming. One alternative is to use existing, high quality, clinical resources as reference standards. To understand the benefits and limitations of this approach, we evaluated the performance of two Agency for Healthcare Research and Quality (AHRQ) Quality Indicators² using existing local databases as reference standards.

METHODS

Using the Clinical Data Repository (CDR), a patient data warehouse for the University of Virginia Health System, we identified two populations of patients (1997-2001): those with wound infections and those who underwent coronary artery bypass surgery (CABG). *Wound infection* and *CABG* were identified by AHRQ Quality Indicator definitions for these conditions. These populations were compared with two existing, high quality databases. The surgical infection control database contains information on infections occurring in patients on the general surgery, trauma, and transplant units.³ The Thoracic-Cardiovascular (TCV) Surgery Outcomes database contains detailed information about patients undergoing TCV surgery, including CABG.

The patient populations identified in the CDR and in the two reference databases were then compared to identify true positives (TP), false positives (FP), and false negatives (FN). (Figure 1) Sensitivity and positive predictive value (PPV) were calculated.

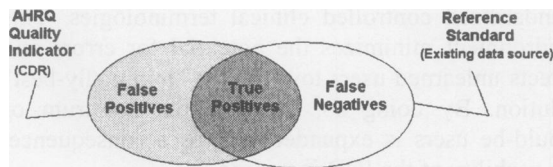


Figure 1. Identifying True/False Positives/Negatives

RESULTS

AHRQ criteria for identifying wound infections did not perform well relative to the reference standard (Table 1); the criteria for identifying CABG utilization performed better, with sensitivity around 0.8 and PPV greater than 0.9. (Table 2)

Year	TP	FP	FN	Sensitivity	PPV
1997	62	117	30	0.674	0.346
1998	61	71	64	0.488	0.462
1999	83	56	48	0.634	0.597
2000	77	45	78	0.497	0.631
2001	40	36	49	0.449	0.526

Table 1. Wound infection (defined by ICD-9-CM) compared to Infection Control database.

Year	TP	FP	FN	Sensitivity	PPV
1997	559	1	87	0.865	0.998
1998	489	12	79	0.861	0.976
1999	494	3	83	0.856	0.994
2000	320	131	79	0.802	0.710
2001	290	105	81	0.782	0.734

Table 2. CABG utilization (defined by DRG) compared to TCV database.

The apparent drop-off in the CABG criteria performance in 2000 and 2001 can be attributed to a change to a new TCV system and subsequent diminished attention to the quality of the original database—our analysis suggests the TCV database is no longer an acceptable reference standard.

CONCLUSION

Existing databases can serve as convenient and inexpensive reference standards for validating and evaluating informatics resources. To be used as reference standards, these databases should meet the following tests:

- Patients/visits must be able to be linked to the test data set.
- The reference database should be of documented high quality and validity. Definitions and entry criteria should be clear.
- Test and reference populations must be truly comparable.

References

1. Hripsak GH, Wilcox A. Reference standards, judges, and comparison subjects. *J Am Med Inform Assoc.* 2002;9:1-15.
2. *AHRQ Quality Indicators.* Nov 2001. Agency for Healthcare Research and Quality, Rockville, MD. <http://www.ahrq.gov/data/hcup/qinext.htm>.
3. Crabtree TD, Pelletier SJ, Raymond DP, et al. Effect of changes in surgical practice on the rate and detection of nosocomial infections: a prospective analysis. (manuscript in review)