Validation of a general practice audit and data extraction tool

David Peiris
Maria Agaliotis
Bindu Patel
Anushka Patel

Background

We assessed how accurately a common general practitioner (GP) audit tool extracts data from two software systems.

Methods

First, pathology test codes were audited at 33 practices covering nine companies. Second, a manual audit of chronic disease data from 200 random patient records at two practices was compared with audit tool data.

Results

Pathology review: all companies assigned correct codes for cholesterol, creatinine and glycated haemoglobin; four companies assigned incorrect codes for albuminuria tests, precluding accurate detection with the audit tool. Case record review: there was strong agreement between the manual audit and the tool for all variables except chronic kidney disease diagnoses, which was due to a tool-related programming error.

Discussion

The audit tool accurately detected most chronic disease data in two GP record systems. The one exception, however, highlights the importance of surveillance systems to promptly identify errors. This will maximise potential for audit tools to improve healthcare quality.

Keywords

cardiovascular diseases; mass screening; clinical coding; data linkage

There is growing interest in the use of primary healthcare data for multiple purposes, including health professional audits, quality improvement programs and research. A critical enabling factor is the availability of data extraction tools to audit patient databases for information such as patient demographics, data quality, disease profiles, risk factor measurements, pathology testing, immunisation and cancer surveillance, use of medicines and Medicare item uptake. Several extraction tools are now available that interface with the major primary healthcare software systems. These tools are used by healthcare providers for practice audits and by many organisations such as Medicare Locals, Aboriginal Community Controlled Health Services (ACCHSs), the Australian Primary Care Collaborative (APCC) program and the National Prescribing Service.[1–3] These tools are also being used for national indicator programs, including the former Divisions of General Practice indicator program,[4] the Queensland Aboriginal and Islander Health Council indicator program,[5] the Northern Territory and the National Key Performance Indicator programs.[6,7] Increasingly, researchers are also using these tools to facilitate data collection.[8] Few studies have assessed the validity of these tools in Australia and generally these have been secondary considerations to the main objectives.[9–11]

In this study, we assessed how accurately one of the most commonly used audit tools extracted data from patient record systems. The study forms part of the Treatment Of cardiovascular Risk in Primary care using Electronic Decision SuppOrt (TORPEDO) study. TORPEDO is a cluster randomised controlled trial examining the effectiveness of a multi-faceted quality improvement intervention to improve cardiovascular disease (CVD) risk screening and management. Sixty health services are participating (40 general practices and 20 ACCHSs) in New South Wales and Queensland. Full details of the study protocol have been published elsewhere.[12]

Methods

TORPEDO involves extracting clinical data for all regularly attending patients (three visits in the previous 2 years and one visit in the previous 6 months) in whom national guidelines recommend CVD risk screening (>35 years if Aboriginal and Torres Strait Islander and >45 years for all others). Participating sites need to be exclusively using either Medical Director™ or Best Practice™ for their electronic health records (EHR) without any use of paper or hybrid paper/electronic recording. These software products currently comprise the majority of EHR systems used in Australia. Extractions are performed at baseline and end of study. Two de-identified data files are generated: (1) an individual patient data file that is sent to the coordinating research institute for trial analyses via a secure file transfer protocol; and (2) an aggregated data file that is sent to a web-based portal via an identical mechanism as that used in the APCC program. Intervention arm sites can perform these extractions monthly and can view peer-ranked performance feedback data.

Variables are extracted on patient demographics, recorded diagnoses of chronic diseases, chronic disease risk factors, pathology tests and medications. We assessed the validity of the data extraction tool in detecting these variables from the two previously described EHR systems via a two-stage process.

Stage 1: Pathology review

For extraction tools to accurately collect pathology data, laboratories must submit test results in a standard format (Health Level Seven [HL-7]) and assign a unique code per test (Logical Observation Identifiers Names and Codes [LOINC]). Prior to validating the tool itself, we reviewed pathology reporting for plasma cholesterol, creatinine and glycated haemoglobin (HbA1c), and urinary albumin to creatinine ratio (UACR). As a condition of participating in TORPEDO we mandated that all sites have their pathology reported in HL-7 format. We then reviewed pathology extracts from a sample of health services covering all the laboratories in the TORPEDO study. These extracts identified what LOINCs were being reported for the tests in question and from these data we determined whether they were concordant with those searched for with the data extraction tool.

Stage 2: Manual case record audit at two general practices

Two hundred records (100 per practice) from patients in the eligible age range for TORPEDO were randomly selected. The two sites were urban Sydney practices participating in the intervention arm of TORPEDO. One site used Medical Director™ and the other used Best Practice™. Audits were conducted as part of a 12-month follow-up visit. Both sites had shown some mild improvement in data quality over the intervention period. A research assistant reviewed the following sections of the record: demographic details, diagnoses, past medical history, physical risk factor measurements, current medications and pathology results, encompassing 23 chronic disease-related variables. Free text entries in progress notes were not reviewed. A fresh data extraction was simultaneously performed and securely sent to the research institute. Data entered from the manual record audit and that obtained from the individual patient data extract were compared. Kappa statistics were used for categorical variables and Bland–Altman plots were constructed for continuous variables with the mean differences and 95% limits of agreement reported.[13] Analyses were conducted using SAS Enterprise Guide (v. 4.5; SAS Institute, Cary, NC, USA)

Results

Pathology audit

Thirty-three sites had pathology data audited covering nine laboratories in New South Wales and Queensland. All laboratories were using the correct LOINC codes for total cholesterol, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, triglycerides, creatinine and HbA1c. There were, however, marked variations in codes used for UACR. Four laboratories used incorrect codes, precluding accurate extraction of results for these tests. All pathology companies were notified and instructed on the correct codes to use for this test.

Manual record audit

Table 1 highlights the patient characteristics for the 200 records audited. *Table 2* outlines the level of agreement/correlation for the 23 variables manually audited. Overall, the majority of variables achieved perfect or near perfect agreement/correlation. There was, however, one notable exception related to chronic kidney disease (CKD). Criteria for CKD were based on national guideline definitions and included a recorded diagnosis of CKD or proteinuria or an estimated glomerular filtration rate (eGFR) < 45ml/min/1.73m^2. There were several cases where the extraction tool seemed

Table 1. Patient characteristics obtained from the manual case record audit tool (n = 200)

Risk factor measurement	Mean (SD) or n (%) in those with available information	No of records with available information (% total)
Demographic information		
Age (years)	66.2 (12.5)	200 (100)
Male	93 (47)	200 (100)
Risk factor measurements		
Current smoker	29 (20)	145 (73)
BMI (kg/m^2)	31.5 (8.8)	97 (49)
Systolic blood pressure (mmHg)	129.1 (15.4)	179 (90)
Diastolic blood pressure (mmHg)	79.1 (10.0)	179 (90)
Pathology laboratory measurements		
Total cholesterol (mmol/L)	5.1 (1.2)	157 (79)
HDL cholesterol (mmol/L)	1.4 (0.4)	155 (78)
LDL cholesterol (mmol/L)	2.8 (0.9)	153 (77)
Triglycerides (mmol/L)	1.9 (1.2)	157 (79)
Urinary albumin to creatinine ratio (mg/mmol)	6.8 (24.3)	93 (47)
Creatinine (umol/L)	78.7 (28.3)	147 (74)
HbA1c (%)	6.4 (1.3)	112 (56)
Past medical history		**n (%)**
Coronary heart disease		30 (15)
Ischaemic stroke/transient ischaemic attack		15 (7.5)
Peripheral vascular disease		6 (3)
Left ventricular hypertrophy		0 (0)
Atrial fibrillation		11 (5.5)
Heart failure		11 (5.5)
Diabetes		41 (20.5)

Abbreviations: BMI = body mass index; HbA1c = glycated haemoglobin; HDL = high-density lipoprotein; LDL = low-density lipoprotein; SD = standard deviation

to inappropriately assign a diagnosis of CKD where no such criteria could be identified from the manual record review. This was subsequently identified to be a problem in the way the tool extracted eGFR values from one pathology company. If the result reported a non-integer value (eg. >90 ml/min), this defaulted to a zero value and hence met one of the CKD criteria (eGFR <45 ml/min). The software company was notified and a fix was put in place. There was also low agreement with UACR tests. This is likely to not be a problem related to the tool per se, but related to incorrect LOINC coding by one of the pathology companies used at that practice.

Discussion

In this study we found that data extraction using one of the most commonly used tools in Australia was highly consistent with manual reviews of the data for all variables except CKD recording which, as a result of this study, was identified to be a programming error and subsequently fixed. We also found wide discrepancies in the coding of CKD-related test results by various laboratories. This affects the ability of any data extraction tool to accurately detect pathology test results for this condition.

An important study limitation is that it was 'fit for purpose' for the chronic disease-related variables used in TORPEDO and for the population recommended for CVD risk screening. We cannot make any assumptions, therefore, about validity for other disease areas. Another limitation is that we looked at interaction between one extraction tool and two GP EHR systems. Although these are the most commonly used systems in Australia, we cannot extrapolate findings to other systems. A final limitation is that we only included data that could potentially be extracted from the EHR. If a practitioner is making free text entries rather than in codable sections of the record then neither the extraction tool nor our manual case record reviewer would have detected those entries.

The Australian health care system is rapidly evolving to incorporate an information communication technology (ICT) infrastructure that will enable data to be shared across multiple platforms. The primary health care sector is a leading driver of this change. GPs and ACCHSs were early adopters of EHRs and are now leading developments in use of data for quality improvement, key performance indicator reporting, secure messaging, and the personally controlled e-health record. The ability to extract data reliably and consistently is central to supporting these activities.

There are two clear recommendations from the study. First, there is a critical need for consistent reporting of pathology tests across all laboratories. Second, the lack of agreement for variables associated with CKD recording, although in retrospect was able to be explained, highlights the importance of establishing regular surveillance procedures that are independent of any validation work conducted by the software vendors themselves. Whenever new code is written, either for the extraction tools or within the EHR system itself, there is potential for errors in the data extraction process. National programs such as the APCC are well placed to establish

Table 2. Agreement/correlation statistics between the audit tool and manual case record audit for 200 patient records from two general practices

Variable	Agreement (kappa) (95% CI where applicable)	
Aboriginal status	1.00	
Smoking status	1.00	
Recorded diagnoses		
Coronary heart disease	0.95 (0.90–1.00)	
Peripheral vascular disease	1.00	
Stroke	0.96 (0.89–1.00)	
Atrial fibrillation	1.00	
Left ventricular hypertrophy	1.00	
Congestive heart failure	0.95 (0.85–1.00	
Diabetes	1.00	
Gestational diabetes	1.00	
Chronic kidney disease	0.24 (0.00–0.48)	
Chronic disease risk factor measurements	**Mean difference* (95% limits of agreement where applicable)**	**No of readings where difference > 2SD (%)†**
Weight (kg)	0.001 (–0.51–0.51)	2 (1.0)
Height (cm)	0	0 (0)
Waist circumference (cm)	0	0 (0)
Systolic blood pressure (mmHg)	0.06 (–1.8–1.6)	3 (1.5)
Diastolic blood pressure (mmHg)	0.14 (-4.9–5.2)	2 (1)
Total cholesterol (mmol/l)	0.02 (–0.37–0.34)	6 (3)
HDL (mmol/l)	0.002 (–0.03–0.04)	0 (0)
LDL (mmol/l)	0.03 (–0.52–0.46)	2 (1)
Triglycerides (mmol/l)	0.02 (–0.65–0.62)	6 (3)
Creatinine (µmol/l)	0.7 (–10.1–8.8)	5 (2.5)
HbA1c (%)	0.001 (–0.18–0.19)	2 (1)
Urinary albumin to creatinine ratio (mg/mmol)	9.7 (–71.7–91.0)	0 (0)

Abbreviations: SD = standard deviation

*The mean difference between the readings obtained from the manual audit vs the audit tool with zero indicating perfect agreement for all readings

†Derived from constructing Bland–Altman plots for each variable

such surveillance systems. This program receives monthly aggregated data from a large number of practices. Automated warning systems for unexpected deviations in these routinely collected data would enable early detection of potential problems. A sentinel network of practices that routinely extracted data could complement such a surveillance system. As we become increasingly reliant on ICT systems for both clinical care and system improvements, it is important that action be taken to ensure these systems are robust.

Implications for general practice

- The audit tool used in this study was reliable for most chronic disease-related measurements with the exception of CKD diagnoses.
- Inaccurate pathology data extracted from audit tools is due to laboratories assigning incorrect codes to test results.
- National initiatives to independently and regularly review the quality of data obtained from data extraction tools are needed.

Authors

David Peiris MBBS, MIPH, PhD, FRACGP, FARGP, Head of Primary Health Care Research, The George Institute for Global Health, University of Sydney, Camperdown, NSW. dpeiris@georgeinstitute.org.au

Maria Agaliotis BSocSci, MPH, Project Officer, Cardiovascular Division, The George Institute for Global Health, University of Sydney, Camperdown, NSW

Bindu Patel BS, MPH, Project Manager, Cardiovascular Division, The George Institute for Global Health, University of Sydney, Camperdown, NSW

Anushka Patel MBBS, SM, PhD, FRACP, FCSANZ, Chief Scientist, The George Institute for Global Health, University of Sydney, Camperdown, NSW.

Ethics approval: The study was approved by the University of Sydney Human Research Ethics Committee (Reference number: 2012/2/2183) and permission was obtained from all participating services to audit records.

Competing interests: None declared.

Provenance and peer review: Not commissioned; externally peer reviewed.

Acknowledgements

We gratefully acknowledge the support of Dr Cedric Meyerowitz and Dr John Hillman in the conduct of this study. David Peiris was supported by an NHMRC Translating Research into Practice Fellowship and now holds a NHMRC Early Career Fellowship. Maria Agaliotis is supported by a NHMRC Senior Research Fellowship.

References

1. Office of Aboriginal and Torres Strait Islander Health. Online Community Health Reporting Environment- OCHREStreams. Available at www.ochrestreams.org.au/pages/welcome.aspx [Accessed 1 April 2013].
2. Knight AW, Caesar C, Ford D, Coughlin A, Frick C. Improving primary care in Australia through the Australian Primary Care Collaboratives Program: a quality improvement report. BMJ Qual Saf 2012;21:948–55.
3. NPS MedicineWise. MedicineInsight. Available at www.nps.org.au/about-us/what-we-do/medicineinsight [Accessed 1 April 2013].
4. Department of Health and Ageing. Divisions of General Practice Program - National Performance Indicators 2009–10. Available at www.phcris.org.au/divisions/reporting/documents/techdetails/NPI_Technical_Details_10_May_2010.pdf [Accessed 1 April 2013].
5. Queensland Aboriginal and Islander Health Council. Core Performance Indicators – Version 2. Available at www.qaihc.com.au/resources/publications/ [Accessed 1 April 2013].
6. Northern Territory Aboriginal Health Forum. NT Aboriginal Health Key Performance Information System. Available at www.nt.gov.au/health/ahkpi/ [Accessed 1 April 2013].
7. Australian Institute of Health and Welfare. Indigenous primary health care key performance indicators. Available at http://meteor.aihw.gov.au/content/index.phtml/itemId/457994 [Accessed 1 April 2013].
8. University of Melbourne. GRHANITE Health Informatics Unit. Available at www.grhanite.com/projects/ [Accessed 1 April 2013].
9. Liljeqvist G, Staff M, Puech M, Blom H, Torvaldsen S. Automated data extraction from general practice records in an Australian setting: Trends in influenza-like illness in sentinel general practices and emergency departments. BMC Public Health 2011;11:435.
10. Schattner P SM, Stanger L, Speak M, Russo K. Clinical data extraction and feedback in general practice: a case study from Australian primary care. Inform Prim Care 2010;18:205–12.
11. Schattner P SM, Stanger L, Speak M, Russo K. Data extraction and feedback - does this lead to change in patient care? Aust Fam Physician 2011;40:623–28.
12. Peiris D, Usherwood T, Panaretto K, et al. The Treatment of cardiovascular Risk in Primary care using Electronic Decision suppOrt (TORPEDO) study: intervention development and protocol for a cluster randomised, controlled trial of an electronic decision support and quality improvement intervention in Australian primary healthcare. BMJ Open 2012;2:6.
13. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;8476:307–10.

correspondence **afp@racgp.org.au**