Informatics in Primary Care 2006;14:235–45 © 2006 PHCSG, British Computer Society

Refereed papers

Smoking status recording in GP electronic records: the unrealised potential

Vanessa Selak MBChB MPH
Research Fellow, Clinical Trials Research Unit, The University of Auckland, New Zealand

Susan Wells MBChB Dip Obs MPH FRNZCGP FAFPHM
National Heart Foundation Research Fellow and Senior Lecturer in Clinical Epidemiology, Section of Epidemiology and Biostatistics, School of Population Health, The University of Auckland, New Zealand

Robyn Whittaker MBChB MPH FAFPHM
Public Health Physician, Health Gain, Funding Team, Waitemata District Health Board, New Zealand

Alistair Stewart BSc DipSc
Senior Research Fellow, Section of Epidemiology and Biostatistics, School of Population Health, The University of Auckland, New Zealand

ABSTRACT

Objective To investigate the recording of smoking status and factors associated with the recording of smoking status in general practitioner (GP) electronic medical records (EMRs) in New Zealand, and the suitability of this source as a prevalence measure.

Setting General practices affiliated with an Auckland-based primary health organisation.

Population Patients registered with 84/107 (78.5%) eligible GPs who had used EMRs for at least a year and had PREDICT-CVD, a web-based cardiovascular disease risk assessment and management decision support program, integrated with their practice software.

Design Audit of EMRs using data from an evaluation of PREDICT-CVD.

Main outcome measures The proportion of EMRs audited (Maori, non-Maori) with smoking status recorded and, among those with smoking status recorded, also Read-coded, and factors associated with greater recording of smoking status.

Results Smoking status was recorded among 49.6% of Maori and 38.3% of non-Maori prior to the installation of PREDICT-CVD. Among those with smoking status recorded, smoking status was also Read-coded among 49.8% of Maori and 62.3% of non-Maori. Factors associated with greater recording of smoking status were installation of PREDICT-CVD, male sex, Maori ethnicity, cardiovascular disease and diabetes. Age was also associated with the recording of smoking status.

Conclusion General practitioner electronic medical records in New Zealand are currently not a suitable source of smoking prevalence data, even if manually searched, as a large proportion of records did not have smoking status recorded. Such records are an even less suitable source of smoking prevalence if data extraction by remote querying (using Read codes) is relied upon. The potential to estimate the prevalence of smoking from GP records has not yet become a reality. Installation of electronic decision support systems, such as PREDICT-CVD, could improve the recording and Read-coding of smoking status, and thereby the availability and accessibility of these data.

Kewords: data collection, information systems, primary health care

Introduction

Tobacco smoking is a major cause of preventable death in New Zealand and a significant contributor to the reduced life expectancy experienced by Maori compared with non-Maori.[1] Accurate and timely smoking prevalence data are required to assess the impact of interventions intended to change smoking habits, inform population-based tobacco control strategies and track trends over time. An alternative to the current sources of smoking prevalence data in New Zealand, such as the New Zealand Census[2] and Household Health Survey,[3] is the use of general practitioner (GP) electronic medical records (EMRs).

The primary care sector is an ideal setting for the identification and documentation of smoking status because 80% of adults visit their GP at least once a year.[3] Identifying patient smoking status increases the rate of clinician intervention, which has been shown to increase smoking cessation.[4] Brief advice to stop smoking by a physician has been reported to increase the absolute rate of smoking cessation by 2.5%.[5] Furthermore, the near complete computerisation of general practices in New Zealand has created an opportunity to obtain, collate and analyse smoking data for populations within the primary care setting. This has occurred within the context of a reorientation of primary health care in New Zealand from a focus on the treatment of individuals and fee-for-service funding, to preventative care for enrolled populations and population-based capitation funding.[6]

Extracting data from EMRs is considered to be a practical and valuable method of providing information about population health characteristics and clinical activity that can be used to improve quality and to monitor healthcare activity and population health needs.[7] This source of data would potentially have the advantage of providing practice and primary health organisation (PHO)-specific data that could be aggregated to a regional/district health board as well as a national level.

Data can be stored as free text or as codes within GP EMRs. Free-text data comprise unstructured notes that typically take a narrative, idiosyncratic form.[8] The analysis of free text is therefore time-consuming because complex searches incorporating multiple words for individual diagnoses and manual checks for accuracy might be required.[9,10] Furthermore, searches of free text are likely to have limited accuracy because not every word used to describe a diagnosis might be included and misspelt words and abbreviations could be missed; therefore manual inspection of at least some, possibly all, records is required.[11]

Codes are shorthand ways of representing health-related concepts.[12] Within medical information systems, codes attempt to standardise the way medical terms are recorded and saved, enabling easy access to, and comparability of, data stored within the EMR through database queries.[13] While a number of coding systems are available internationally, Read codes are the most widely used coding system by New Zealand GPs.[14] Named after their creator, English GP James Read,[12] Read codes were developed in the early 1980s in order to assist with the capture of diagnoses in GP EMRs.[9] In general, Read codes are arranged hierarchically, so that at each level data are more detailed.[12]

We investigated the recording of smoking status and factors associated with the recording of smoking status in GP EMRs in New Zealand, and the suitability of this source as a prevalence measure for Maori and non-Maori.

Methods

This study used data from an evaluation of PREDICT-CVD, a web-based, electronic decision support program for cardiovascular disease (CVD) risk assessment and management in primary care that enables GPs to assess CVD risk and provide individualised evidence-based recommendations for patient management.[15] At the time of the evaluation, most of the data used by PREDICT-CVD had to be entered by GPs because self-population from data already within the EMR was limited to age, sex, ethnicity and fasting blood test results. PREDICT-CVD has been fully integrated with the electronic patient management system (PMS) MedTech-32^{TM} (which uses 5-byte Read codes) in the Auckland-based PHO ProCare Health Ltd under the brand name 'Prompt', and is currently being used by more than 300 GPs.

The PREDICT-CVD Evaluation Study was a before-and-after study conducted within general practices affiliated with ProCare Health Ltd. Eligible GPs were current members of ProCare Health Ltd who had used MedTech-32^{TM} for EMRs and for receiving electronic laboratory results for at least one year prior to April 2004, who had PREDICT-CVD installed in the practice, and who had registered patients. Patients whose notes were included in the audit had been seen by an eligible and consenting GP in their practice during a four-week period, one month after the date of first use of PREDICT-CVD (post-PREDICT-CVD) and/or in the same four-week period 12 months previously (pre-PREDICT-CVD). The target patient groups for CVD risk assessment according to the New Zealand CVD risk management guidelines are: Maori, Pacific and Indian subcontinent men aged $\geq$ 35 years and women aged $\geq$ 45 years and all other men aged $\geq$ 45 years and women aged $\geq$ 55 years.[16] The EMRs of all

Maori patients meeting these criteria were included to maximise the number of Maori participants. The EMRs of a random selection (using the random numbers table function in Microsoft Excel) of 15% of non-Maori patients meeting these criteria for each eligible, consenting GP for each time period were included. A detailed description of the PREDICT-CVD Evaluation Study, which focuses on an evaluation of the effectiveness of PREDICT-CVD, has been submitted for publication.

Among other variables, the audit investigated the recording of smoking status: whether the patient was a current smoker, non-smoker or past smoker (not smoking for > 12 months). This involved a manual inspection of each EMR with the audit nurses instructed to inspect components of the records in the following prescribed systematic order: first classification, then front page, then history, then screening, then inbox, then daily record, then outbox and then specialist letters. Smoking status as well as the first location in which it had been found (if recorded) were documented. The location of recorded smoking status was used to identify whether or not smoking status had been Read-coded. Within MedTech-32^{TM}, smoking status is systematically recorded and saved as a Read code in classification, history and screening, but not in other sections of the EMR. Audit nurses sought additional information from other sites in the medical record to clarify ambiguous or conflicting information.

Descriptive and stratified analysis was undertaken using EpiInfo (Version 3.2.2). Multivariate analysis was undertaken using SAS (Version 9.1). Odds ratios were adjusted for the GP (the primary sampling unit) and practice as well as patient characteristics, including age group, sex, ethnicity (Maori or non-Maori), the presence of existing CVD or diabetes, and holding a High Use Health Card (government subsidy for those with medical conditions requiring frequent GP visits) or Community Services Card (government subsidy for lower income families). A mixed logistic regression model was used, in which GPs were regarded as random effects and all other variables were regarded as fixed effects.

Results

Of the 107 eligible GPs, 84 (78.5%) consented to participate; 18 declined; four were on leave and one could not be contacted, reason unknown. Compared to non-participators, GPs who participated were similar in terms of age, sex and mean number of years since graduation. However, those that did not participate were more likely to have had PREDICT-CVD installed between August 2003 and May 2004 than the previous year. A total of 3564 audits were conducted; 1680 for the pre-PREDICT-CVD period (August 2001 to June 2003) and 1884 for the post-PREDICT-CVD period (August 2002 to June 2004).

Table 1 describes the demographic characteristics of the patients seen in the two time periods of interest. With the exception of previous CVD (chi-sq 6.74; $P = 0.009$) there were no differences between these two groups in terms of age group, sex, ethnicity, having diagnosed diabetes or holding a High Use Health Card or a Community Services Card.

Recording of smoking status and Read-coding of recorded smoking status

Smoking status was recorded among 49.6% of Maori and 38.3% of non-Maori prior to the installation of PREDICT-CVD. The recording of smoking status was modestly greater after the installation of PREDICT-CVD (59.3% of Maori and 47.9% of non-Maori). Among those with smoking status recorded, smoking status was also Read-coded among 49.8% of Maori and 62.3% of non-Maori prior to the installation of PREDICT-CVD. The Read-coding of recorded smoking status was also modestly greater after the installation of PREDICT-CVD (51.2% of Maori and 67.0% of non-Maori; see Table 2).

Factors associated with the documentation of smoking status

Factors associated with greater recording of smoking status were installation of PREDICT-CVD (OR 1.60; 95% CI 1.4–1.9), male sex (1.27; 1.1–1.5), Maori ethnicity (1.78; 1.4–2.2), history of CVD (1.44; 1.2–1.8) and diagnosed diabetes (2.40; 1.9–3.0) (see Table 3). Age was also associated with the recording of smoking status ($P < 0.0001$). Smoking status was recorded with increasing frequency as age increased to those aged 55–64 years and then it decreased with age.

Factors associated with greater Read-coding of smoking status, among EMRs with smoking status recorded, were installation of PREDICT-CVD (OR 1.41; 95% CI 1.1–1.8) and diagnosed diabetes (1.63; 1.2–2.2). Age had an effect on Read-coding ($P = 0.04$) and it appears that this effect came from the people aged 35–44 with recorded smoking status being less likely to be Read-coded compared with older people.

Approximately 5% of all records were re-audited by a different audit nurse to assess the quality of data collection. There was a 17.7% discrepancy in the recording of smoking status between audits, but this

Table 1 Characteristics of audited populations before and after PREDICT-CVD installation

Variable	Pre-PREDICT-CVD ($n = 1680$)		Post-PREDICT-CVD ($n = 1884$)	
	n	%	*n*	%
Age group (years)				
> 85	73	4.3	85	4.5
75–84	269	16.0	354	18.8
65–74	377	22.5	420	22.3
55–64	488	29.1	535	28.4
45–54	366	21.8	372	19.7
35–44	107	6.3	118	6.3
Sex				
Male	853	50.8	914	48.5
Female	827	49.2	970	51.5
Ethnicity				
Maori	474	28.2	484	25.7
Non-Maori	1206	71.8	1400	74.3
High Use Health Card status				
HUHC	161	9.6	190	10.1
No HUHC	1519	90.4	1694	89.9
Community Services Card status				
CSC	761	45.3	824	43.7
No CSC	919	54.7	1060	56.3
Diagnosed diabetes or on diabetes treatment				
Diabetes	247	14.7	297	15.8
No diabetes	1433	85.3	1587	84.2
Previous CVD event or on nitrates				
CVD	327	19.5	434	23.0
No CVD	1353	80.5	1450	77.0

discrepancy reduced to 4.3% when smoking status had been Read-coded.

Discussion

Key findings

There is the potential to estimate the prevalence of smoking from GP EMRs in New Zealand, but this potential has not yet become a reality according to this study. In the pre-PREDICT-CVD sample, only 38–50% of patients had smoking status recorded. While the recording of smoking status was greater in the post-PREDICT-CVD sample (48–59%), the magnitude of this difference was only modest. Therefore, GP EMRs are currently not a suitable source of smoking prevalence data, even if manually searched. GP EMRs are an even less suitable source of smoking prevalence data if data extraction by remote querying is relied upon, given that, among records with smoking status recorded, smoking status was also Read-coded in only 50–62% (pre-PREDICT-CVD) and 51–67% (post-PREDICT-CVD).

While all patients in this study met the criteria for CVD risk assessment, the recording of smoking status varied according to patient characteristics. That men and Maori were more likely to have smoking status recorded could be due to the known higher prevalence of smoking in these population groups. Increased recording of smoking status among people with CVD and diabetes reflects good clinical practice because of the importance of smoking cessation in these high-risk groups. Although smoking prevalence is known to decline with age, there is no reason why smoking status should be recorded less often in older people. Stopping smoking at any age confers health benefits.

Table 2 Recording and Read-coding of smoking status by ethnicity

Variable	*n*	(*N*)	%
Pre-PREDICT-CVD			
Smoking status recorded			
Maori	235	(474)	49.6
Non-Maori	462	(1206)	38.3
Smoking status Read-coded (among EMRs with smoking status recorded)			
Maori	117	(235)	49.8
Non-Maori	288	(462)	62.3
Post-PREDICT-CVD			
Smoking status recorded			
Maori	287	(484)	59.3
Non-Maori	670	(1400)	47.9
Smoking status Read-coded (among EMRs with smoking status recorded)			
Maori	147	(287)	51.2
Non-Maori	449	(670)	67.0

(*N*), total number.

The discrepancy of recorded smoking status between the original and repeat audits reflects the complexity of retrieving data from uncoded clinical records.

Study validity

Smoking status was only able to be measured in this study if it had been documented in the EMR. It was not possible to ascertain the accuracy of the classification of smoking status from this study because an alternative source of data (for example, separate patient notes, patient interviews/questionnaires) was not obtained. One study found that agreement between records of smoking status and patient-completed questionnaires was only moderate ($\kappa = 0.50$)[17] and that 46% of patients who reported themselves as ex-smokers were misclassified as being never smokers by their GPs.[17]

There was potential for evidence of Read-coding to have been missed due to the order of manual inspection. However, as only the 2% of records with smoking status recorded on the front page could have been affected by this, it is highly unlikely that our final results would be altered.

The recording of smoking status and the Read-coding of smoking status (among those with smoking status recorded) are independent measures and were analysed separately. Factors that might have been associated with the documentation of smoking status were therefore able to be assessed for their effects on the recording of smoking status and the Read-coding of recorded smoking status separately.

Comparison with other studies

GPs were members of the same PHO (distributed widely across the Auckland region), had the same PMS, and were adopters of an electronic decision support system. Therefore, they might be different from other GPs in New Zealand and potentially provide an overestimate of the general level of primary care recording of smoking status.

However, these findings are consistent with those of other studies conducted on morbidity data in GP EMRs in New Zealand.[9,14,18] Findings from similar UK studies indicate that UK GPs are recording and coding smoking status more than GPs in New Zealand.[7,19]

Conclusions

Despite the priority given to smoking and the need to record smoking status according to national guidelines, smoking status was often not recorded and even when recorded was often not Read-coded. There is an increasing need for accurate and useful information to enable the reorientation of primary health care from

Table 3 Recording of smoking status and Read-coding of recorded smoking status

Variable	Recording of smoking status Adjusted odds ratio (95% CI)	Read-coding of recorded smoking status (among EMRs with smoking status recorded) Adjusted odds ratio (95% CI)
PREDICT-CVD		
Post-PREDICT-CVD	1.60 (1.4–1.9)*	1.41 (1.1–1.8)*
Pre-PREDICT-CVD	1.00	1.00
Age group (years)		
> 85	0.34 (0.2–0.5)*	0.82 (0.4–1.9)
75–84	0.65 (0.5–0.8)*	1.19 (0.8–1.8)
65–74	1.00	1.00
55–64	1.68 (1.4–2.1)*	1.18 (0.8–1.7)
45–54	1.39 (1.1–1.8)*	1.07 (0.7–1.6)
35–44	0.73 (0.5–1.1)	0.46 (0.2–0.9)*
Sex		
Male	1.27 (1.1–1.5)*	1.18 (0.9–1.6)
Female	1.00	1.00
Ethnicity		
Maori	1.78 (1.4–2.2)*	0.89 (0.6–1.2)
Non-Maori	1.00	1.00
High Use Health Card status		
HUHC	1.13 (0.9–1.5)	0.82 (0.5–1.2)
No HUHC	1.00	1.00
Community Services Card status		
CSC	1.12 (0.9–1.3)	1.09 (0.8–1.5)
No CSC	1.00	1.00
Diagnosed diabetes or on diabetes treatment		
Diabetes	2.40 (1.9–3.0)*	1.63 (1.2–2.2)*
No diabetes	1.00	1.00
Previous CVD event or on nitrates		
CVD	1.44 (1.2–1.8)*	0.80 (0.6–1.1)
No CVD	1.00	1.00

Mixed logistic regression model included GP, practice, PREDICT-CVD, age group, sex, ethnicity, Community Services and High Use Health Card status, diabetes and CVD.
GPs were regarded as random effects and all other variables were regarded as fixed effects.
* $P < 0.05$

a focus on the treatment of individuals to a greater consideration of the health of populations.[6] Such information is required to support and inform needs assessments, service planning, funding, delivery and monitoring, the co-ordination of provider activities and patient care, the improvement of continuity of care and clinical decision making, and quality improvement processes.[6]

Improving the recording and systematic coding of smoking status in primary care requires a commitment to an information culture. Primary health services need to foster an information-rich environment, use data to inform planning/service provision and to feed back data, analyses and quality measures to improve patient care.

ACKNOWLEDGEMENTS

We wish to thank all the GPs from ProCare Health Ltd for participating in the study, their practice teams who made us welcome and the audit nurses and staff from the Diabetes Project Trust.

REFERENCES

1 Laugesen M and Clements M. *Cigarette Smoking Mortality Among Maori, 1954–2028*. Wellington: Te Puni Kokiri; 1998.
2 Statistics New Zealand. *Introduction to the Census*. Wellington: Statistics New Zealand; 2001.
3 Ministry of Health. *A Portrait of Health: key results of the 2002/03 New Zealand Health Survey*. Wellington: Ministry of Health; 2004.
4 Fiore MC, Bailey WC, Cohen SJ *et al. Treating Tobacco Use and Dependence: clinical practice guideline*. Rockville, MD: US Department of Health and Human Services. Public Health Service; 2000.
5 Silagy C and Stead LF. Physician advice for smoking cessation. *The Cochrane Database of Systematic Reviews* 2004.
6 King A. *The Primary Health Care Strategy*. Wellington: Ministry of Health; 2001.
7 Horsfield P and Teasdale S. Generating information from electronic patient records in general practice: a description of clinical care and gender inequalities in coronary heart disease using data from over two million patient records. *Informatics in Primary Care* 2003;11(3):137–44.
8 AHIMA e-HIM Work Group on Computer-Assisted Coding. Delving into computer-assisted coding. *Journal of the American Health Information Management Association* 2004;75(10):48A–H.
9 Hall J, Tomlin A, Martin I and Tilyard MW. A general practice minimum data set for New Zealand. *New Zealand Medical Journal* 2002;115(1163). www.nzma.org.nz/journal/115–1163/200/content.pdf
10 Hall J and Martin I. The Dunedin RNZCGP Research Unit Computer Research Network: an update to 2003 and beyond. *New Zealand Family Physician* 2003;30(3):181–5.
11 Britt H, Beaton N and Miller G. General practice medical records: why code? Why classify? *Australian Family Physician* 1995;24:612–15.
12 Booth N. What are the Read Codes? *Health Libraries Review* 1994;11(3):177–82.
13 Jones NF. *Monitoring Health Using Routinely Recorded General Practice Clinical Data: evaluation of an Auckland surveillance project*. Thesis. Auckland: University of Auckland; 2001.
14 Didham R, Martin I, Wood R and Harrison K. Information technology systems in general practice medicine in New Zealand. *New Zealand Medical Journal* 2004; 117(1198). www.nzma.org.nz/journal/117–1198/977/content.pdf
15 Wells S and Jackson R. Online management of cardiovascular risk in New Zealand with PREDICT: getting evidence to the 'moment of care'. *Health Care & Informatics Review Online* 2005;March.
16 New Zealand Guidelines Group. *The Assessment and Management of Cardiovascular Risk*. Wellington: New Zealand Guidelines Group; 2003.
17 Mant J, Murphy M, Rose P and Vessey M. The accuracy of general practitioner records of smoking and alcohol use: comparison with patient questionnaires. *Journal of Public Health Medicine* 2000;22(2):198–201.
18 Rafter N, Connor J, Hall J *et al*. Cardiovascular medications in primary care: treatment gaps and targeting by absolute risk. *New Zealand Medical Journal* 2005; 118(1223):U1676.
19 Hippisley-Cox J, O'Hanlon S and Coupland C. Association of deprivation, ethnicity, and sex with quality indicators for diabetes: population-based survey of 53 000 patients in primary care. *British Medical Journal* 2004;329:1267–70.

FUNDING

This study was supported by funds from Waitemata District Health Board and The Future Forum. Vanessa Selak received a training endowment from the New Zealand Population Health Charitable Trust to undertake this research. Sue Wells is the recipient of a National Heart Foundation Research Fellowship.

ETHICAL APPROVAL

The PREDICT-CVD Evaluation Study was approved by the Auckland Regional Ethics Committee (AKY/04/07/185).

CONFLICTS OF INTEREST

None.

ADDRESS FOR CORRESPONDENCE

Vanessa Selak
Clinical Trials Research Unit
The University of Auckland
Private Bag 92019
Auckland Mail Centre
Auckland 1142
New Zealand
Tel: +64 9 373 7999
Fax: +64 9 373 1710
Email: v.selak@ctru.auckland.ac.nz

Accepted November 2006