Informatics in Primary Care 2011;19:241–50

Refereed paper

Using a data entry clerk to improve data quality in primary care electronic medical records: a pilot study

Michelle Greiver MD MSc CCFP
Director, North Toronto Research Network, Toronto, Canada and Assistant Professor, Department of Family and Community Medicine, University of Toronto, Canada

Jan Barnsley BSc MES PhD
Associate Professor, Department of Health Policy, Management & Evaluation, University of Toronto, Canada and Adjunct Scientist, Institute for Clinical Evaluative Sciences, Toronto, Canada

Babak Aliarzadeh MD MPH
Data Manager, North Toronto Research Network, Toronto, Canada

Paul Krueger BSc MHSc MSc PhD
Associate Professor, Associate Director Research Program, Department of Family and Community Medicine, University of Toronto, Canada

Rahim Moineddin MSc PhD
Associate Professor, Department of Family and Community Medicine, University of Toronto, Canada, Scientist, Institute for Clinical Evaluative Sciences, Toronto, Canada and Associate Professor, Dalla Lana School of Public Health, University of Toronto, Canada

Debra A Butt MD MSc CCFP FCFP
Assistant Professor, Department of Family and Community Medicine, University of Toronto, Canada

Edita Dolabchian MSc
Research Associate, North Toronto Research Network, Toronto, Canada

Liisa Jaakkimainen MD MSc
Scientist, Institute for Clinical Evaluative Sciences, Institute for Clinical Evaluative Sciences, Toronto, Canada and Associate Professor, Department of Family and Community Medicine, University of Toronto, Canada

Karim Keshavjee MD MBA CCFP CPHIMS
Adjunct Assistant Professor, University of Victoria, Victoria, Canada

David White MD CCFP
Chief, Department of Family & Community Medicine, North York General Hospital, North York, Ontario, Canada and Associate Professor, Family and Community Medicine, University of Toronto, Canada

David Kaplan MSc MD CCFP
Associate Family Physician-in-Chief, Chair of the Research Ethics Board, North York General Hospital, North York, Ontraio, Canada and Assistant Professor, Department of Family and Community Medicine, University of Toronto, Canada

For the North Toronto Research Network (NorTReN)

ABSTRACT

Background The quality of electronic medical record (EMR) data is known to be problematic; research on improving these data is needed.

Objective The primary objective was to explore the impact of using a data entry clerk to improve data quality in primary care EMRs. The secondary objective was to evaluate the feasibility of implementing this intervention.

Methods We used a before and after design for this pilot study. The participants were 13 community based family physicians and four allied health professionals in Toronto, Canada. Using queries programmed by a data manager, a data clerk was tasked with re-entering EMR information as coded or structured data for chronic obstructive pulmonary disease (COPD), smoking, specialist designations and interprofessional encounter headers. We measured data quality before and three to six months after the intervention. We evaluated feasibility by measuring acceptability to clinicians and workload for the clerk.

Results After the intervention, coded COPD entries increased by 38% ($P = 0.0001$, 95% CI 23 to 51%); identifiable data on smoking categories increased by 27% ($P = 0.0001$, 95% CI 26 to 29%); referrals with specialist designations increased by 20% ($P = 0.0001$, 95% CI 16 to 22%); and identifiable interprofessional headers increased by 10% ($P = 0.45$, 95 CI –3 to 23%). Overall, the intervention was rated as being at least moderately useful and moderately usable. The data entry clerk spent 127 hours restructuring data for 11 729 patients.

Conclusions Utilising a data manager for queries and a data clerk to re-enter data led to improvements in EMR data quality. Clinicians found this approach to be acceptable.

Keywords: computerised/standards, data collection/standards, data quality, health care/methods, medical records systems, primary care, quality assurance

What is known

- Electronic medical record (EMR) data quality is known to be problematic.
- There are few interventional studies addressing this problem; interventions have generally led to modest improvements.

What this paper adds

- Data queries programmed by a data manager followed by EMR data entry by a data clerk led to large increases in structured data over a short period for a chronic disease (chronic obstructive pulmonary disease), identifiable smoking categories and specialist designations.
- There was no significant increase in interprofessional encounter designations, a change that relied on modification of clinician behaviour.
- The acceptability of the intervention to clinicians and the cost indicate that larger studies of similar interventions are feasible.

Introduction

Data quality matters: 'you cannot improve what you cannot measure'.[1] The transition from paper records to electronic medical records (EMRs) has led to expectations that electronic healthcare data collected as part of routine practice will be available for quality improvement activities, surveillance, research and chronic disease management.[2–12]

The quality of the information collected fundamentally depends on the quality and integrity of data entered in the charts. However, problems with the data include inconsistent or missing diagnostic coding and risk factor designation, 'dirty data' (misspelled words, inconsistent word strings, free text strings instead of structured data), missing 'meta-data' (referral to 'Dr Smith', where physician specialty is not listed) and data entered in inconsistent or incorrect database fields.[13–21] During the transition to EMRs, training is often focused on using and entering data in individual patient records, with limited emphasis on consistent data entry and future auditing capabilities. Family physicians and their practice teams may not be aware of the importance of this issue[22] and have many competing demands on their time and resources.[23] Once the EMR transition is complete, physicians may have limited time, incentives or tools to modify and

improve data that were initially entered as unstructured free text, added to fields not meant for these specific data or entered in several different areas of the EMR.

Systematic reviews of data quality have noted many descriptive studies but few interventional studies designed to improve data quality in primary care EMRs.[24,25] Most interventional studies used education or individualised feedback.[25]

Based on the existing literature and our clinical experience, the underlying ideas for this study were: (1) data entry difficulties were common during the transition to EMRs; (2) problems were subsequently not systematically corrected or managed; (3) a data manager may be able to identify some problematic areas; (4) trained data entry clerks could efficiently re-enter data; and (5) once the initial data entry is done, practices may be able to maintain reasonable data quality using tools such as data manuals.

The primary purpose of this study was to explore the impact of an intervention designed to improve data quality in the EMRs of community based family physicians. The secondary purpose was to evaluate the feasibility of implementing this intervention.

Methods

Study design

We used a before and after design. We first used professionally programmed data queries to measure data quality and identify gaps. The intervention consisted of assigning data re-entry away from healthcare providers: we used a data entry clerk for this work. We then re-used the original queries after the intervention so that the change could be calculated.

Participants

We recruited community based family physicians in Toronto, Canada who were members of an interdisciplinary primary care organisation (the North York Family Health Team) and were using the Nightingale On Demand® EMR. Forty-three family physicians in the family health team used this software. We recruited a convenience sample of 13 physicians that have used EMRs for at least two years (to ensure that early transition efforts were completed), as indicated by the presence of EMR-based progress notes for over two years. We also recruited four allied health providers who had provided clinical services to patients registered to participating physicians during the study. Eligible patients included all active patients registered with the practices who were age 18 or more at the time of the audit.

Intervention

The data queries were programmed by the North Toronto Research Network (NorTReN) data manager; NorTReN is one of 10 practice-based research networks currently participating in the Canadian Primary Care Sentinel Surveillance Network (CPCSSN), Canada's first multidisease primary care electronic record surveillance system. A local data manager oversees EMR data collection, cleaning and transmission to the central data repository for each network.[13]

We examined the change in data quality in four areas of the EMR: diagnostic coding for a chronic health condition (chronic obstructive pulmonary disease or COPD), structured categories for a risk factor (smoking), structured specialist referral designation (meta-data) and interprofessional encounter designation. The rationale for selecting these four areas is that CPCSSN data managers have found that health conditions are not consistently coded in the patient health profile, smoking status is recorded using a large number of free text terms, and specialist referral designations are not consistently available.[13] Interdisciplinary care provision is not currently collected for CPCSSN, but is important for primary care system planning.

In order to improve the generalisability of the findings, we used data queries and extraction tools available within the practices through their EMR interface. That is, the data manager did not use queries that required direct access to the underlying EMR databases, as this method would not be available to practices wishing to repeatedly query their own EMR for data quality improvement purposes.

A research associate used the programmed queries to audit the EMR and to record baseline measures. A data clerk was then tasked with re-entering the data, as follows: (1) with physician permission, adding the International Classification of Diseases ninth revision (ICD9) code 496 for COPD (the most common code for this condition in the CPCSSN database) in the patient health profile when free text indicating COPD was found; (2) duplicating free text smoking data using a drop down list classifying a patient as a current smoker, ex-smoker or never smoked; (3) adding referral designations to all specialists in the master referral list that comply with College of Physicians and Surgeons of Ontario specialist designation; and (4) adding standardised interprofessional encounter headers to the EMR as a drop down list if these were not previously present and informing allied health professionals. The clerk was trained by the research associate and was given data manuals with screen shots. The clerk entered ten training records for each of the four areas which were audited by the research associate for accuracy. After the initial audits, clinicians were given data manuals with suggested methods of data entry (see

Appendix A available online at: www.radcliffepublishing.com/journals/J12_Informatics_in_primary_care/supplementary%20papers.htm).

Outcome measures

Data quality measures were: (1) COPD designations in the patient health profile that were coded using ICD9; (2) patient records that had data on tobacco use in a structured format; (3) specialty referrals within a three-month period with structured specialist designation; and (4) encounters by allied health professionals indicated as interprofessional care within a three-month period.

We extracted EMR data at baseline and at three to six months after the intervention.

We evaluated feasibility through acceptability to clinicians and by measuring time and cost for the data clerk. Clinicians (family physicians and allied health providers) were given questionnaires incorporating usefulness (perception of the degree that the process would enhance job performance) and usability (perception of the degree that the process would be free from effort).[26,27] The questionnaires are shown in Appendix B available online at: www.radcliffepublishing.com/journals/J12_Informatics_in_primary_care/supplementary%20papers.htm. The clerk submitted hours worked to the research associate; we recorded total amount of time (including training) for each data aspect.

Analysis

The significance of the change in the proportion of each measure of quality was assessed using McNemar's test for paired samples. We used descriptive and summary statistics for physician and practice characteristics and for acceptability to clinicians. All tests were two-sided and P-values < 0.05 were considered statistically significant. Data were analysed using SAS 9.2 (SAS Institute).

This study was approved by the North York General Hospital's Research Ethics Board. All physicians and allied health professionals who participated in the study provided signed, informed consent.

Results

Physician and practice characteristics

Physician and practice characteristics are shown in Table 1. All physicians were in group practices. There were three office locations but physicians shared a single EMR server, accessed from a remote location. There were 11 729 eligible patients at the time of the initial audit in September 2010, and 11 554 patients at the time of the second audit in March 2011.

A summary of the changes in data structure before and after the intervention is shown in Table 2. The

Table 1 Physician and practice characteristics*

Physician characteristics		N=13
Female	*N* (%)	10 (77)
Age	Median (range)	36 (34–59)
CCFP	*N* (%)	12 (92)
Years since graduation from medical school	Median (range)	11 (6–34)
Canadian medical school graduate/foreign medical school graduate	*N/n*	12/1
Number of physicians at the practice location	Median (range)	5 (5–7)
Number of nurses at the practice location	Median (range)	1.4 (0.7–2.0)
Duration of EMR use	Median (range)	4 (2–7)
Number of patients registered to the physician	Median (range)	800 (660–1388)
Number of patients seen in an average week	Median (range)	80 (48–120)
Number of hours providing office-based patient care per week	Median (range)	25 (15–45)

Note: CCFP = Certificate of the College of Family Physicians of Canada.
* Obtained from self reports at study entry; based on full time equivalent.

Table 2 Coded or structured data present in the EMR

Data element	Baseline (%)	Post intervention (%)	Difference*: % (*P*, 95% CI)
Coded COPD entries: number coded/total number with COPD in heath profile (%)	44/75 (59)	102/106 (96)	38 (*P*=0.0001, 23–51)
Structured smoking categories: number with structured data/total number with smoking data (%)	4,285/6039 (71)	6208/6317 (98)	27 (*P*=0.0001, 26–29)
Specialist designations in referral letters: number of structured designations/total number of specialist referrals (%)	831/1619 (51)	1177/1649 (71)	20 (*P*=0.0001, 16–22)
Interprofessional encounter headers: number of audited charts with appropriate headers/total number of audited charts (%)	25/89 (28)	42/111 (38)	10 (*P*=0.45, –3–23)

Note: CI = confidence interval; COPD = chronic obstructive pulmonary disease.
* Difference may not be exact due to rounding. In a three month period. A 10% sample of interprofessional encounters for each of four allied health providers was randomly audited.

proportion of coded or structured data elements increased for all categories studied, although this was not statistically significant for interprofessional encounter headers.

Coded COPD entries in the patient health profile

Prior to the intervention, 59% of COPD entries in the patient health profile were numerically coded using ICD9. ICD9 codes for this disease at baseline were 496, 492 and 491. The clerk entered all new codes as ICD9 496.

The total number of COPD patients increased during the project because physicians concurrently verified patients with COPD as part of a quality assurance project in which they were given a list of patients who were potential COPD candidates because they were aged 45 years or over, non-asthmatic and used medications indicated for COPD (tiotropium, salbutamol, inhaled steroids).[28] Those verified by physicians as having COPD were entered and coded by the clerk as ICD9 496.

Data were re-audited in March 2011. The percentage of coded COPD entries increased to 96%.

Pick list data on tobacco risk category

Data about tobacco use were audited in October 2010. After the audit, the clerk accessed charts where free text tobacco information had been entered in the patient health profile and added data using a structured drop down list (current smoker, ex-smoker, never smoked).

A follow-up audit occurred in March 2011. At baseline, 51% of patients had information on tobacco usage in their health profile, compared with 55% of patients during the second audit. Of those with tobacco data present, 71% had structured data on smoking; this was usually a checkbox indicating either smoker or non-smoker. After the intervention, 98% of patients with data on smoking had identifiable categories. Current smokers were not identifiable using standard EMR queries prior to the intervention. After the intervention, 732 patients (12% of those with data on smoking status) could be identified as current smokers.

Structured specialist referral designations

We audited the charts for a three-month period prior to the intervention (27 June 2010 to 27 September 2010). The data clerk added specialist designations to the master referral list in October 2010. We re-audited the charts for a three-month period following the intervention (15 November 2010 to 15 February 2011). One physician went on maternity leave between the first and second audits, and her data were censored from both audits. Identifiable specialist designations increased from 51 to 71%.

Standardised interprofessional encounter designations

It was not possible to use the EMR software to audit encounters for the presence of interprofessional headers. EMR logs were used to identify all encounters done by an allied health provider for the three-month period prior to the intervention (27 June 2010 to 27 September 2010) and for a three-month period after the intervention (15 November 2010 to 15 February 2011). One allied health provider from each of the four categories in the family health team was randomly chosen, and then a randomly chosen 10% sample of the health provider's encounters was manually audited. Usage varied by allied health provider role; the nurse almost never used the headers (2% of encounters prior to the intervention, 0% after), whereas the dietitian started using headers routinely (from 29% of encounters to 90% of encounters). The social worker and clinical pharmacist were using headers for all encounters prior to the intervention, and this did not change. Interprofessional headers increased by 10%; the change was not statistically significant.

Participant ratings of interventions

Participating clinicians rated the usability and usefulness of this approach. Results are shown in Table 3 for usability and Table 4 for usefulness. Eleven of the 13 eligible physicians returned the questionnaires: one physician did not respond and one was on maternity leave. All four allied health professionals responded. There were three questions dealing with usability and four questions dealing with usefulness.

Data entry clerk workload

Including training, the data entry clerk spent 3 hours recoding COPD, 53 hours restructuring the smoking data and 70 hours adding specialist designations to the master list. Interprofessional headers were added in less than 1 hour, for a total of 127 hours spent on all activities at a cost of $1905, or $147 per physician.

COPD coding and tobacco categories required chart by chart data entry. The master referral list was shared between all practices as part of the common server, and therefore data were entered once for all practices. Interprofessional encounter headers were shared between all providers within an office but not between offices, and were therefore replicated three times for the three office locations studied.

Discussion

Principal findings

A data manager can program queries to discover data quality issues and a trained data entry clerk can rapidly re-enter uncoded and unstructured data in the EMR as coded, structured and consistent data for groups of practices. We found significant increases in coded or

Table 3 Participant rating of usability of intervention*

Category	Rating: number of responses (% of all responses† for each category††)				
	Not at all usable	Not very usable	Neutral	Moderately usable	Very usable
COPD coding	0 (0)	0 (0)	0 (0)	16 (48)	17 (52)
Smoking category restructuring	0 (0)	0 (0)	1 (3)	8 (24)	24 (73)
Specialist designation	0 (0)	0 (0)	4 (12)	14 (42)	15 (45)
Interprofessional encounter headers	0 (0)	0 (0)	0 (0)	3 (25)	9 (75)

* Obtained from self reported perceptions of usability and usefulness28 at study exit.
† Percentage may not add up to 100 due to rounding. †† From 11 physicians and 4 allied health professionals; there were three questions for each category, for a total of 33 responses from physicians and 12 responses from allied health professionals.

Table 4: Participant rating of usefulness of intervention*

Category	Rating: number of responses (% of all responses† for each category††)				
	Not at all useful	Not very useful	Neutral	Moderately useful	Very useful
COPD coding	0 (0)	4 (9)	4 (9)	9 (20)	27 (61)
Smoking category restructuring	0 (0)	4 (9)	9 (20)	7 (16)	24 (55)
Specialist designation	0 (0)	2 (5)	16 (36)	13 (30)	12 (27)
Interprofessional encounter headers	2 (13)	0 (0)	3 (19)	9 (56)	2 (13)

* Obtained from self reported perceptions of usability and usefulness28 at study exit.
† Percentage may not add up to 100 due to rounding. †† From 11 physicians and 4 allied health professionals; there were four questions for each category, for a total of 44 responses from physicians and 16 responses for allied health professionals.

structured data with this intervention for three of the four areas we studied. There was good acceptance by providers, and the time spent by the data clerk did not seem excessive. These findings indicate that a larger study would be feasible.

Implications of the findings

Uncoded data were present in all four areas we studied for these community based family physicians. Standardising data elements may assist in developing comparisons over time within practices, between individual providers, between groups of physicians and between different jurisdictions.[21]

An unexpected finding was that a list of current smokers could not be generated using EMR queries prior to the intervention, due to the variability in data entry. This was possible after data entry. The ability to identify a population at risk can enable EMR features associated with improved quality of care such as chart alerts or recalls.[29–31]

We believe that the magnitude of the changes is likely clinically important. A change of 5% or more has been used to determine the minimal clinically important difference[32–34] and we report larger changes in this study.

Comparison with the literature

Difficulties with coding have been reported previously.[14,16,19,20,35] Other studies have found that the use of data extraction was possible; similar to our study, practices required external support for this.[36]

Studies on data improvement in primary care EMRs have relied on audit, feedback and training, with moderate effects.[37–41] In this study, we report larger changes through the involvement of non-clinicians in data management and data re-entry in EMRs. The area that required a change in clinician data entry behaviour (interprofessional encounter headers) did not change significantly.

Limitations

Limitations for this pilot study include lack of population diversity and a single EMR system. However, the fact that the study took place in a community based primary care setting indicates the potential to conduct a larger community based trial that would be more broadly generalisable.

There was a concurrent quality assurance effort for COPD. This increased the total number of patients defined as having this condition (denominator) after the intervention and may have affected the results for this aspect of data quality.

The information is limited to the EMR application we studied. However, we believe that similar restructuring can be undertaken with other EMR software applications commonly used in primary care; data issues have been found with every EMR studied as part of CPCSSN.[13,19]

The cost and workload of the data management and data entry clerk, as well as the amount of training

required will vary according to the EMR type, the structure of the EMR used, provider data entry habits and availability of technical resources.

In this study, we did not validate the accuracy of data or address data completeness. In Toronto, 18.2% of persons age 12 years or over are current smokers[42] compared to 12% of adults with smoking data identified in our study, so it is possible that there was missing or misidentified data on current smokers in the EMR.

We tested the overall effect of the implementation of our intervention. We did not capture impact on provider behaviour, such as changes in data entry habits, for three of the four measures.

Members of primary care practices continually enter data, and may reverse some of the improvements over time. We measured data over a relatively brief interval and longer studies will be needed to quantify the loss of data quality. However, the physician's generally positive ratings of usefulness suggest that improvements in data quality provide a recognisable benefit to EMR users, who may thus improve their own data entry. Interventions such as repeated audit and feedback, as well as ongoing maintenance activities (such as the use of data clerks at fixed intervals) would be needed to assess and maintain data quality.

Further research and recommendations

Additional studies and methods could include measures of data entry reliability such as re-audits of data entry samples; measures of validity (comparison with reference standard); sustainability of the changes; qualitative methods to explore perceptions and barriers to this approach; impact on provider behaviour such as improved quality and consistency of data entry; and an economic analysis of the cost of data entry clerks in various primary care settings and EMR applications.

Collaboration with EMR vendors to improve the structure of their underlying databases may be worthwhile. Efforts are underway to implement EMR data content standards.[43,44] Vendors could increase the amount of structured data that can be captured as part of clinical care in their applications, could improve queries and could automate data linkages. For example, allied health professionals could automatically have searchable designations linked to encounters.

Conclusions

In this study, the use of a data entry clerk led to fairly large and rapid improvements in EMR data quality. We found increases in COPD coding, standardised tobacco risk categories and structured specialist designation, with reasonable rates of clinician acceptance and workload for the clerk.

ACKNOWLEDGEMENTS

This study was funded by the Primary Health Care System Program; we acknowledge the support of the Ontario Ministry of Health and Long-Term Care. The views expressed do not necessarily reflect those of the Ministry. The Canadian Primary Care Sentinel Surveillance Network is funded by the Public Health Agency of Canada. The views expressed herein do not necessarily represent the views of the Public Health Agency of Canada.

REFERENCES

1 Newcomer LN. Perspective: physician, measure thyself. *Health Affairs* 1998;17(4):32–5.
2 Romanow R. *Building on Values: the future of health care in Canada*. Ottawa: National Library of Canada, 2002.
3 Garcia Rodriguez LA and Perez Gutthann S. Use of the UK General Practice Research Database for pharmacoepidemiology. *British Journal of Clinical Pharmacology* 1998;45(5):419–25.
4 Bartholomeeusen S, Kim CY, Mertens R, Faes C and Buntinx F. The denominator in general practice, a new approach from the Intego database. *Family Practice* 2005;22(4):442–7.
5 Biermans MC, Spreeuwenberg P, Verheij RA, de Bakker DH, de Vries Robbe PF and Zielhuis GA. Striking trends in the incidence of health problems in The Netherlands (2002–05). Findings from a new strategy for surveillance in general practice. *European Journal of Public Health* 2009;19(3):290–6.
6 Hueston WJ, Mainous III AG, Ornstein S, Pan Q and Jenkins R. Antibiotics for upper respiratory tract infections: follow-up utilization and antibiotic use. *Archives of Family Medicine* 1999;8(5):426–30.
7 Hall J, Tomlin A, Martin I and Tilyard M. A general practice minimum data set for New Zealand. *New Zealand Medical Journal* 2002;115(1163):U200.
8 Stocks NP, McElroy H, Sayer GP and Duszynski K. Acute bronchitis in Australian general practice. A prescription too far? *Australian Family Physician* 2004; 33(1–2):91–3.
9 Lagasse R, Desmet M, Jamoulle M *et al. European Situation of Routine Medical Data Collection and their Utilization for Health Monitoring: Euro-Med-Data Final Report*. Universite Libre de Bruxelles: Brussels, 2001.
10 Smeeth L, Thomas SL, Hall AJ, Hubbard R, Farrington P and Vallance P. Risk of myocardial infarction and stroke after acute infection or vaccination. *New England Journal of Medicine* 2004;351(25):2611–18.
11 Schade R, Andersohn F, Suissa S, Haverkamp W and Garbe E. Dopamine agonists and the risk of cardiac-valve regurgitation. *New England Journal of Medicine* 2007;356(1):29–38.

12 Have paper records passed their expiry date? *Canadian Medical Association Journal* 2005;173(7):725.
13 Birtwhistle R, Keshavjee K, Lambert-Lanning A *et al.* Building a pan-Canadian primary care sentinel surveillance network: initial development and moving forward. *Journal of the American Board of Family Medicine* 2009; 22(4):412–22.
14 Jordan K, Porcheret M and Croft P. Quality of morbidity coding in general practice computerized medical records: a systematic review. *Family Practice* 2004;21(4): 396–412.
15 Whitelaw FG, Nevin SL, Milne RM, Taylor RJ, Taylor MW and Watt AH. Completeness and accuracy of morbidity and repeat prescribing records held on general practice computers in Scotland. *British Journal of General Practice* 1996;46(404):181–6.
16 Selak V, Wells S, Whittaker R and Stewart A. Smoking status recording in GP electronic records: the unrealised potential. *Informatics in Primary Care* 2006;14(4):235–41; discussion 242.
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 Hollowell J. The General Practice Research Database: quality of morbidity data. *Population Trends* 1997(87): 36–40.
19 Mukhi S and Keshavjee K. *Developing a Primary Care Electronic Medical Record Chronic Disease Surveillance System in Canada: data quality and lessons learned.* Canadian Association of Health Services and Policy Research: Calgary, 2009.
20 Birtwhistle R, Keshavjee K, Martin K and Lambert-Lanning A. *Improving Data Quality in EMRs for Chronic Disease Management.* Family Medicine Forum: Calgary, 2009.
21 Canadian Institutes for Health Information. *CPCSSN Data Analysis Report, Cycle 2A.* Canadian Institutes for Health Information: Toronto, 2010.
22 Audet AM, Doty MM, Shamasdin J and Schoenbaum SC. Measure, learn, and improve: physicians' involvement in quality improvement. *Health Affairs (Millwood)* 2005;24(3):843–53.
23 Yarnall KS, Pollak KI, Ostbye T, Krause KM and Michener JL. Primary care: is there enough time for prevention? *American Journal of Public Health* 2003; 93(4):635–41.
24 Thiru K, Hassey A and Sullivan F. Systematic review of scope and quality of electronic patient record data in primary care. *BMJ* 2003;326(7398):1070.
25 Brouwer H, Bindels P and Weert H. Data quality improvement in general practice. *Family Practice* 2006; 23(5):529–36.
26 Bagozzi RP, Davis FD and Warshaw PR. Development and test of a theory of technological learning and usage. *Human Relations* 1992;45(7):659–86.
27 Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 1989;13(3):319–39.
28 Qaseem A, Wilt TJ, Weinberger SE *et al.* Diagnosis and management of stable chronic obstructive pulmonary disease: a clinical practice guideline update from the American College of Physicians, American College of Chest Physicians, American Thoracic Society, and European Respiratory Society. *Annals of Internal Medicine* 2011;155(3):179–91.
29 Szilagyi PG, Bordley C, Vann JC *et al.* Effect of patient reminder/recall interventions on immunization rates: a review. *Journal of the American Medical Association* 2000; 284(14):1820–7.
30 Kaczorowski J, Goldberg O and Mai V. Pay-for-performance incentives for preventive care. *Canadian Family Physician* 2011;57(6):690–6.
31 Holbrook A, Thabane L, Keshavjee K *et al.* Investigators ftCI: individualized electronic decision support and reminders to improve diabetes care in the community: COMPETE II randomized trial. *Canadian Medical Association Journal* 2009;181(1–2):37–44.
32 Copay AG, Subach BR, Glassman SD, Polly DW Jr and Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine Journal* 2007;7(5):541–6.
33 Jaeschke R, Singer J and Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Controlled Clinical Trials* 1989;10(4): 407–15.
34 Greiver M, Kasperski J, Barnsley J and Rachlis V. *Measuring Quality Improvements in Preventive Care Services in the First Two Family Health Networks in the Greater Toronto Area.* Ontario College of Family Physicians: Toronto, 2006.
35 Khan NF, Harrison SE and Rose PW. Validity of diagnostic coding within the General Practice Research Database: a systematic review. *British Journal of General Practice* 2010;60(572):e128–36.
36 Schattner P, Saunders M, Stanger L, Speak M and Russo K. Clinical data extraction and feedback in general practice: a case study from Australian primary care. *Informatics in Primary Care* 2011;18(3):205–12.
37 Linder JA, Rigotti NA, Schneider LI *et al.* An electronic health record-based intervention to improve tobacco treatment in primary care: a cluster-randomized controlled trial. *Archives of Internal Medicine* 2009;169(8): 781–7.
38 de Lusignan S, Stephens PN, Adal N and Majeed A. Does feedback improve the quality of computerized medical records in primary care? *Journal of the American Medical Informatics Association* 2002;9(4):395–401.
39 Porcheret M, Hughes R, Evans D *et al* and the North Staffordshire General Practice Research Network. Data quality of general practice electronic health records: the impact of a program of assessments, feedback, and training. *Journal of the American Medical Informatics Association* 2004; 11(1):78–86.
40 Lopez-Picazo Ferrer JJ, Agullo Roca F, Villaescusa Pedemonte M and Cerezo Corbalan JM. Clinical data that are essential for the primary care clinical records: an experience of evaluation and improvement. *Aten Primaria* 2002;30(2):92–8.
41 Brown PJ and Warmington V. Info-tsunami: surviving the storm with data quality probes. *Informatics in Primary Care* 2003;11(4):229–33; discussion 234.
42 Table 105–0502 – Health indicator profile, two year period estimates, by age group and sex, Canada, prov-

inces, territories, health regions (2007 boundaries) and peer groups, occasional. Ottawa: Statistics Canada, 2010.
43 Sullivan-Taylor P, Flanagan T, Harrison T and Webster G. Development of a draft pan-Canadian primary health care electronic medical record content standard. *Studies in Health Technology and Informatics* 2011;164:385–91.
44 Sullivan-Taylor P, Mukhi S, Martin-Rhee M and Webster G. Data that makes a difference in quality improvements in primary health care: approaches through a pan-Canadian voluntary electronic medical record source. *Studies in Health Technology and Informatics* 2011;164:367–71.

ADDRESS FOR CORRESPONDENCE

Michelle Greiver
North Toronto Research Network
240 Duncan Mill Road Suite 705
Toronto
Ontario M3B 3S6
Canada
Email: mgreiver@rogers.com

Accepted January 2012

Supplementary material is available online at: www.radcliffepublishing.com/journals/J12_Informatics_in_primary_care/supplementary%20papers.htm