The Comparison of Models of Primary Care in Ontario (COMP-PC) study: methodology of a multi-faceted cross-sectional practice-based study

SIMONE DAHROUGE, WILLIAM HOGG, GRANT RUSSELL, ROBERT GENEAU, ELIZABETH KRISTJANSSON, LAURA MULDOON, SHARON JOHNSTON

ABSTRACT

Background: Many industrialized nations have initiated reforms in the organization and delivery of primary care. In Ontario, Canada, salaried and capitation models have been introduced in an attempt to address the deficiencies of the traditional fee-for-service model. The Ontario setting therefore provides an opportunity to compare these funding models within a region that is largely homogeneous with respect to other factors that influence care delivery. We sought to compare the performance of the models across a broad array of dimensions and to understand the underlying practice factors associated with superior performance. We report on the methodology grounding this work.

Methods: Between 2004 and 2006 we conducted a cross-sectional mixed-methods study of the fee-for-service model, including family health groups, family health networks, community health centres and health service organizations. The study was guided by a conceptual framework for primary care organizations. Performance across a large number of primary care attributes was evaluated through surveys and chart abstractions. Nested case studies generated qualitative provider and patient data from 2 sites per model along with insights from key informants and policy-makers familiar with all models.

Results: The study recruited 137 practices. We conducted 363 provider surveys and 5361 patient surveys, and we performed 4108 chart audits. We also conducted interviews with 40 family physicians, 6 nurse practitioners, 24 patients and 8 decision-makers. The practice recruitment rate was 45%; it was lowest in fee-for-service practices (23%) and in family health networks (37%). A comparison with all Ontario practices in these models using health administrative data demonstrated that our sample was adequately representative. The patient participation (82%) and survey scale completion (93%) rates were high.

Conclusions: This article details our approach to performing a comprehensive evaluation of primary care models and may be a useful resource for researchers interested in primary care evaluation.

Simone Dahrouge is the manager of research operations at the C.T. Lamont Primary Health Care Research Centre, the research arm of the Department of Family Medicine, Faculty of Medicine at the University of Ottawa, Ottawa, Ontario, Canada. **William Hogg** is a professor at the Department of Family Medicine, the Department of Epidemiology and Community Medicine, and the Institute of Population Health, University of Ottawa, and director of the C.T. Lamont Primary Health Care Research Centre. **Grant Russell** is an associate professor at the Department of Family Medicine and the Department of Epidemiology and Community Medicine at the University of Ottawa, and a clinician investigator at the C.T. Lamont Primary Health Care Research Centre. **Robert Geneau** is an adjunct professor in the Department of Family Medicine at the University of Ottawa and principal scientist at the C.T. Lamont Primary Health Care Research Centre. **Elizabeth Kristjansson** is a principal scientist at the Institute of Population Health, University of Ottawa. **Laura Muldoon** is a lecturer in the Department of Family Medicine at the University of Ottawa and a principal scientist at the C.T. Lamont Primary Health Care Research Centre. **Sharon Johnston** is an assistant professor at the Department of Family Medicine and a principal scientist at the C.T. Lamont Primary Health Care Research Centre, University of Ottawa.

Competing interests: None declared.

Funding source: Funding for this research was provided by the Primary Health Care Transition Fund of Ontario Ministry of Health and Long-Term Care. The views expressed in this report are the views of the authors and do not necessarily reflect those of the Ontario Ministry of Health and Long-Term Care.

Correspondence: Simone Dahrouge, Manager, Research Operations, C.T. Lamont Primary Health Care Research Centre, Élisabeth-Bruyère Research Institute, 43 Bruyère St., Ottawa ON K1N 5C8; sdahroug@scohs.on.ca

As a growing body of evidence reveals the importance of primary care to the health of populations, there is increasing interest in the efficient, effective and equitable delivery of these services. In response, many industrialized nations have initiated reforms in the organization and delivery of primary care with the aim of optimizing care delivery.[1] Primary care is funded in several different ways by different countries. Capitation funding provides a fixed annual sum to a practice for the care of each patient registered with that practice. Fee-for-service funding provides payment to a practice according to services delivered, such as patient consultations and type of care delivered. In a salaried service, the health care providers are employed and practice income is not dependent on the number of services provided or the number of patients served. Recently some countries have made efforts to introduce quality- or performance-related payments into existing payment structures.[2-4] There is little evidence to indicate which models of funding of primary care deliver better services, and international comparisons are difficult to interpret because differences are not confined to funding models.

The situation in Ontario, Canada, provides an excellent opportunity to compare funding models for primary care because the 3 major models described above have been used side by side in recent years. This enables comparisons to be made largely unconfounded by differences in gross domestic product, percent spending on health care, patient characteristics and professional training. Over the past 2 decades, Ontario has developed an array of diverse models of primary care delivery but little information on their comparative performance is available to guide further reform initiatives. In 2002, the government of Canada established the Primary Health Care Transition Fund, an $800-million commitment to help provinces and territories develop and sustain new approaches to primary health care delivery. In this article we report on the methodology of a mixed-methods practice-based study sponsored from this fund, the Comparison of Models of Primary Care in Ontario (COMP-PC). We studied fee-for-service (FFS) practices (including the traditional FFS model and reformed family health group model), a capitation-based system called health service organizations (HSOs), a model of multidisciplinary community health centres (CHCs) employing salaried physicians with a focus on community needs, and a relatively new model of physician-run group practices, the family health networks (FHNs), which incorporated extended-hour coverage, financial support for information technology and a blended remuneration formula of capitation, performance bonuses and fee for service.

Our aim was to measure the impact of funding models of primary care on patient self-reported quality of care and on provider adherence to recommended standards of care. In this article we detail the study design and the methods used for data collection. We describe how we categorized and sampled practices using different funding models, how we collected information on processes of care that might explain model differences and how we measured the outcomes of quality and adherence. This large study used a complex methodology that cannot be sufficiently described in associated articles. This article, therefore, serves as an elaboration of the methods that will be reported in a succinct form elsewhere.

Methods

Objectives. The objectives of the COMP-PC study were to describe 4 funding models (FFS, HSOs, CHCs and FHNs), to measure and compare the quality of primary care delivered and to better understand aspects of practice organization that may influence the health care experience of patients and the quality of care they receive. The process and outcome evaluation were theory based[5] and guided by a conceptual framework (Fig. 1).[6]

Design. The COMP-PC project was a cross-sectional mixed-methods study of primary care practices involving quantitative data collection and a nested qualitative case study using a subset of 2 sites per model. The Ottawa Hospital Research Ethics Board approved the study. Figure 2 summarizes the study sampling approach and eligibility criteria.

Study population. The study involved primary care practices, their providers and patients. We also interviewed key informants and policy-makers who had in-depth knowledge of each model.

Sample size. The study measured the performance of primary care practices across numerous outcomes. Because we expected the measure of performance in disease prevention to require the greatest number of measurements, it was used to estimate sample size. Performance in disease prevention was measured as the adherence to recommended guidelines for 6 manoeuvres (see Table 1, section 2.2). A patient's disease prevention score was the proportion of manoeuvres performed to manoeuvres for which he or she was eligible.

Sample size was calculated using a minimum clinically important difference of 0.5 standard deviation, with an alpha value of 0.05 and a beta value of 0.20, and was chosen to control for the family-wise error rate and variance of the cluster (cluster correlation coefficient of 0.2).[7] The basic unit of random selection was the prac-

See related article: Hogg W, Dahrouge S, Russell G, Tuna M, Geneau R, Muldoon L, Health promotion activity in primary care: performance of models and associated factors. *Open Med* 2009;3(3):165-173

tice. The recommendation that resulted from this calculation was to include data from 40 practices per model and data from at least 30 patients per practice. Owing to budgetary and time limitations, the number of practices was reduced to 35. We aimed to collect up to 50 surveys at each practice (instead of 30) to compensate for the possibility that surveys would not be adequately completed.

For the nested case study, we selected 8 practices (2 per model) from within the sites recruited for the cross-sectional study to allow for methodological and data triangulation.We stopped conducting interviews after wereached an acceptable level of data saturation for each model and for each category of respondent (providers, patients and key informants).

Study participants: practices

Eligibility. For practical reasons, we excluded practices in the far north of the province.[1] Over the course of the recruitment period, we noted that the majority of practices under the traditional FFS model had converted to family health groups (FHGs), a modified FFS model introduced as the study was getting underway. At the time of recruitment, the main difference between the FHG and the traditional FFS models was that FHG practices were required to register their patients and provide extended hours of service, for which they received additional compensation.[8] Three months before the end of recruitment a decision was made to include FHG practices within the traditional FFS group, and we endeavoured to enrol those FFS practices previously deemed non-eligible because they had converted to FHGs. In this document we refer to both models as FFS.

Consent to participate was required from at least half of the physicians and nurse practitioners in the organization. Practices were also required to have operated under their model for at least 1 year and provide general primary care services. Practices also provided consent to allow the study investigators to access the information related to their practice contained in health administrative databases housed at the Institute for Clinical Evaluative Sciences (ICES). Practices were considered a group if the individual providers shared at least 4 of the following 5 items or resources: office space, staff, expenses, patient records and on-call duties. Practices with different geographic locations (addresses) were considered separate even if they were linked in a network.[9]

Sampling strategies and recruitment. All of the CHC, HSO and FHN practices in Ontario and a randomly selected group of 197 FFS–FHG practices were invited to participate. Forty-two of these FFS–FHG practices were found to be not eligible, leaving 155 eligible FFS–FHG practices. For the nested case study, we used a typical case sampling strategy to select the sites.[10] Practice sites were invited to participate in this qualitative component if they typified the model to which they belonged in size and composition. Practices needed to be large enough to allow sufficient provider interviews to permit data saturation within that model. We recruited 1 urban and 1 rural practice from each model, with the exception of HSOs; 2 urban sites were selected for HSOs because these organizations are concentrated in urban areas. The sample base covered practices serving approximately 90% of the provincial population of 12.6 million at the time of sampling.

Study invitation materials were mailed to eligible practices. Follow-up was done through a combination of mailings, telephone calls and face-to-face visits. We also sought the support of the model's central organizational structure where one existed (i.e., CHCs and HSOs) in delivering study information and promoting participation.

Sites were offered C$2000 in recognition of the time required by professionals and administrative staff to participate in the study. An additional C$500 was paid to those practices participating in the qualitative component of the study. Recruitment and data collection took place from June 2005 to June 2006.

Study participants: providers

Eligibility. Physicians and nurse practitioners working at the practice were eligible to participate in the study if they had practised at that site for at least 1 year or 6 months, respectively; the participating site was the principal site of their clinical practice; the majority of their services were devoted to primary care; and the majority of their patients were over the age of 17 years.

Sampling strategies and recruitment. Practices were asked to invite all eligible providers to participate in the study and were informed that participation by at least half of the eligible providers was required for the practice to be included in the study; 363 providers participated. Practices electing to also participate in the qualitative component provided names of family physicians and nurse practitioners who were interested in interviews. For 2 sites with multiple providers, this process yielded only 2 providers. In these cases, snowball sampling was then used to recruit providers through the first contact.

Study participants: patients

Eligibility. Patients were eligible to complete the survey if they were patients of consenting providers, 18 years of age or older, not severely ill or cognitively impaired, not known to the survey administrator and able to communicate in English or French either directly or

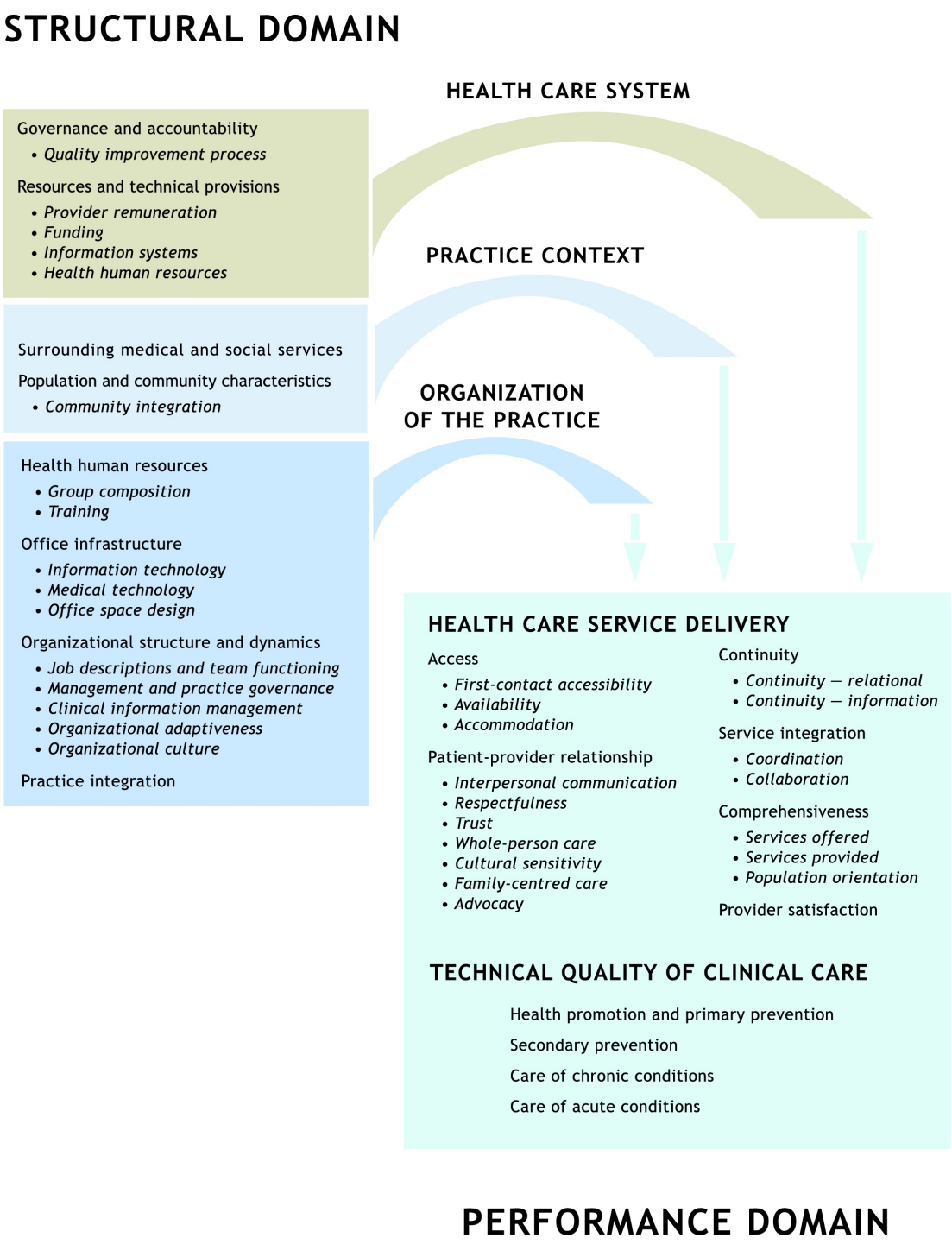

Figure 1: Conceptual framework for primary care organizations. Adapted from Hogg W, Rowan M, Russell G, Geneau R, Muldoon, L. A conceptual framework for primary care: the importance of a structural domain. *International Journal for Quality in Health Care* 2008;20(5):308-313. Used with permission of the journal and Oxford University Press.

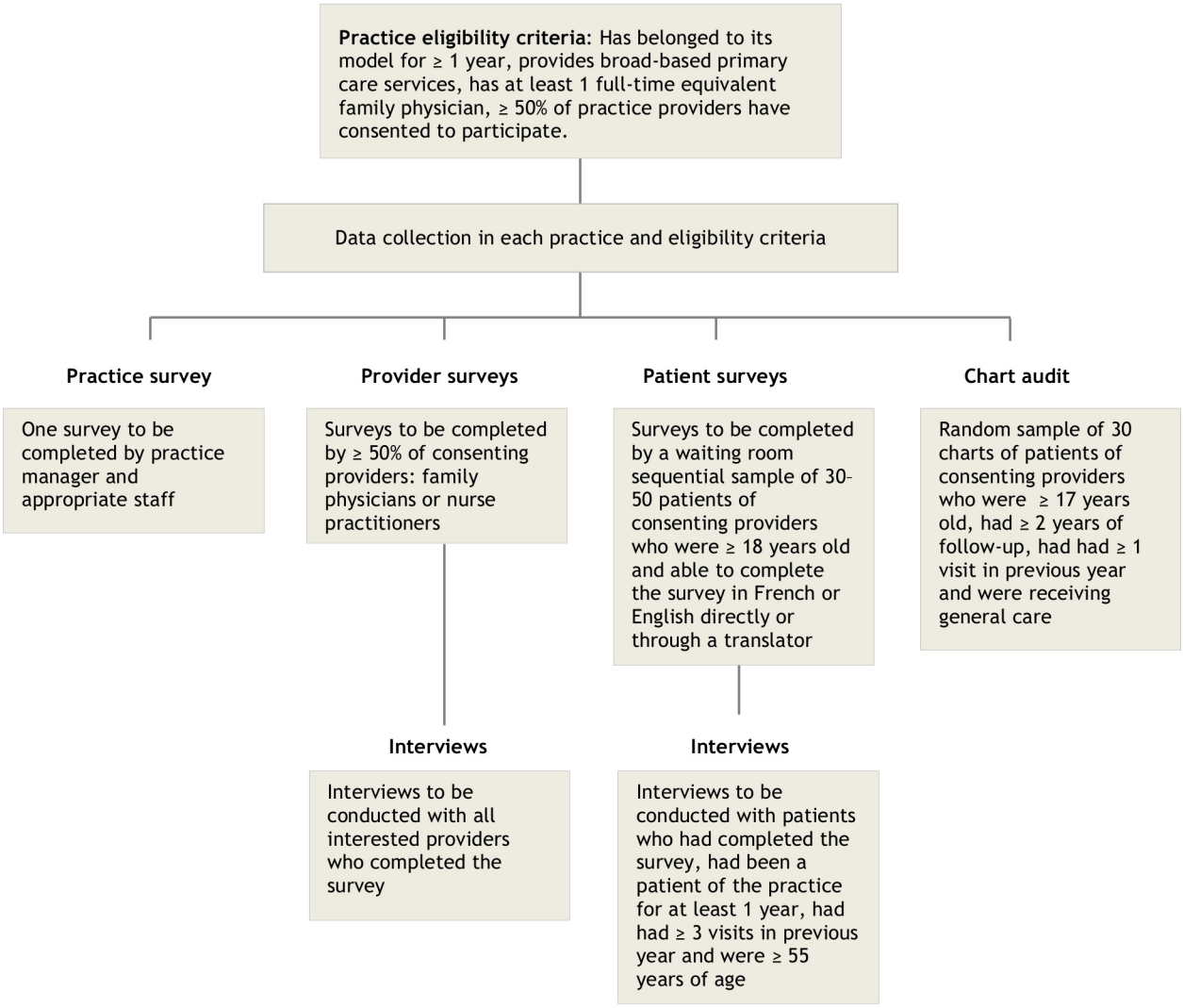

Figure 2: Practice-based study recruitement and eligibiity flow chart. Two sites per model (a total of 8 practices) were selected for the in-depth qualitative assessment (the interview phase). In these practices, we interviewed a total of 40 family physicians, 6 nurse practitioners and 24 patients.

Table 1: Scales and indicators

Attribute		Data source Pr	Pt	C	Scales and indicators (source)
STRUCTURAL DOMAIN — ORGANIZATION					
1	**Health care system**				
1.1	Governance and accountability				Qualitative evaluation
1.2	Resources and technical provisions	X			Provider remuneration (IH)
2	**Practice context**				
2.1	Surrounding medical and social services	X			Types of and distance to institutions and professionals servicing the community (IH)
		X	X	X	Patient sociodemographic and economic information (adapted from PCAT) and patient health status (PCAS)
2.2	Population and community characteristics	X			Team structure and sociodemographic characteristics (IH)
3	**Organization of the practice**				
3.1	Health human resources	X			Provider sociodemographic information (IH) Practice staffing and structure (IH)
3.2	Office infrastructure	X			Information technologies implemented (adapted from National Physician Survey)[1]
3.3	Organizational structure and dynamics	X			Duration of operation as a practice and within its model (IH) Roles and responsibilities at the site and extent of sharing resources (adapted from National Physician Survey)[1] Duration of regular visit, hours of operation, provider payment structure, revenues and operation costs (IH) Quality control audits
				X	Chart organization (IH)
PERFORMANCE DOMAIN — OUTCOME					
1	**Quality of health care service delivery**				
1.1	Access	X	X		First contact accessibility scales (PCAT)
			X		First contact utilization scale (PCAT) Duration of today's consultation (IH)
		X			Practice accepting new patients (IH) Duration of routine visit (IH) Hours of operation and on-call hours (IH) Disability access (IH)
1.2	Patient–provider relationship		X		Humanism scale[2] Trust scale[3] Wait time in clinic (IH)
		X	X		Cultural competency scales (PCAT) Family centredness scales (PCAT)

Attribute		Data source Pr	Pt	C	Scales and indicators (source)
1.3	Continuity		X		Ongoing care scale (PCAT) Relationship with practice and provider (IH)
1.4	Service integration	X			Coordination scale (PCAT) Coordination: information system scale (PCAT) Extent of sharing resources (adapted from National Physician Survey)[1]
1.5	Comprehensiveness	X			Services offered (PCAT) Community orientation scale: reach out (PCAT) Community orientation scale: needs assessment (PCAT) Community orientation scale: monitor (PCAT)
		X	X		Community orientation scale (PCAT)
			X		Reason for visit (IH)
1.6	Provider satisfaction	X			Provider satisfaction scale (National Physician Survey)[1] Remuneration preferences (IH)
2	**Technical quality of clinical care delivery**				
2.1	Chronic disease management			X	Manoeuvres performed in adherence with recommended guidelines: • Coronary artery disease: aspirin, beta blocker, statins • Diabetes: HbA_1C test frequency, angiotensin-converting enzyme inhibitor or angiotension receptor blocker, seen by an ophthalmologist or optometrist, feet checked or patient referred to a chiropodist or podiatrist • Congestive heart failure: angiotensin-converting enzyme inhibitor or angiotension receptor blocker, beta blocker Intermediate clinical outcomes: • Hypertension: blood pressure results • Diabetes: HbA_1C result
2.2	Disease prevention			X	Manoeuvres performed in adherence with recommended guidelines: • High risk for influenza: influenza vaccine • 50 years of age or older: colorectal cancer screening by sigmoidoscopy or hemoccult stool test • Females 50—69 years of age: breast cancer screening by mammography and clinical examination • Females under 60 years of age: cervical screening • 65 years of age or older: clinical hearing examination • 65 years of age or older: screening for visual impairment
2.3	Health promotion		X		Subjects discussed at that visit in adherence with recommended guidelines: • Healthy foods and unhealthy foods • Home safety, such as getting and checking smoke detectors and storing medicines safely • Family conflicts • Exercise • Check and discussion of the medications the patient is taking • Tobacco, smoking • Alcohol consumption • Fall prevention

Attribute		Data source Pr	Pt	C	Scales and indicators (source)
2.4	Management of acute conditions			X	Appropriate prescription of antibiotic in adherence with recommended guidelines: • Sore throat • Urinary tract infection

IH = questions developed in house, PCAT = Primary Care Assessment Tool, PCAS = Primary Care Assessment Survey.

This table shows the indicators and scales used in the quantitative evaluation of the models. Assessment of the various attributes of the structural domain was principally informed by the qualitative evaluation. The performance domain was measured through surveys and chart abstractions (Pr = practice or provider survey, Pt = patient survey, C = chart). Most health service delivery attributes were measured using the PCAT scales. The instruments were supplemented with questions designed specifically for this study. The technical quality of care delivery was assessed by comparing the extent to which the care delivered was consistent with recommended guidelines for the management of patients in primary care.

1. College of Family Physicians of Canada. 2002. Updated data release of the 2001 National Family Physician Workforce Survey. www.cfpc.ca/local/files/Programs /Janus%20project/NFPWS2001_Final_Data_Release_rev_en pdf. [Accessed 7 Mar 2007]
2. Hauck FR, Zyzanski SJ, Alemagno SA, Medalie JH. Patient perceptions of humanism in physicians: effects on positive health behaviors. *Fam Med* 1990;22(6):447–452.
3. Hall MA, Zheng B, Dugan E, Camacho F, Kidd KE, Mishra A, et al. Measuring patients' trust in their primary care providers. *Med Care Res Rev* 2002;59(3):293–318.

Table 2: Study participation rate by model

Variable	CHC	HSO	FHN	FFS	Overall
Practices					
Approached, no.	53	69	104	197	423
Eligible, no.	51	65	94	155	365
Participated, no.	35	32	35	35	137
Response rate, %	69	49	37	23	45
Providers					
Participated,* no.	182	42	81	58	363
Patients					
Eligible, no.	1591	1590	1583	1758	6522
Participated, no.	1219	1273	1494	1375	5361
Response rate, %	77	80	94	78	82
Chart abstraction	1050	958	1050	1050	4108

CHC = community health centre, HSO = health service organization, FHN = family health network, FFS = traditional fee-for-service and family health group.

* Provider recruitment was relinquished to the practice manager. We did not track the actual participation rate other than to ensure it was at least 50%.

through a translator. Patients participating in the qualitative component of the study were also required to have been patients of the practice for at least 1 year and to have attended at least 3 appointments. We gave preference to those 55 years of age or older.

Sampling strategies and recruitment. Following a prepared script, receptionists introduced the study and handed an invitation letter to all patients presenting for their appointment on the day of survey administration. Using another prepared script, the survey administrator provided more detailed information about the study, verified whether the patient met the full set of eligibility criteria and invited eligible patients to participate. In practices participating in the qualitative component of the study, the survey administrator invited patients who had completed surveys to take part in an in-depth interview at a later date, until 6–8 agreed.

Chart audit

Eligibility. Chart abstraction was limited to the charts of regular patients of consenting care providers who were 17 years of age or older at the time of their last visit and had at least 2 years of information, with at least 1 visit in the previous year. Patients were excluded if they had died or had left the practice in the previous 2 years, had used the practice for specialized services only (e.g., foot care), were known to the chart abstractor or were staff members of the practice.

Random selection. In practices with paper-based charting, the total length of the shelves containing the charts was divided into 60 "similar distance" sections, and the fifth chart from the start of each section was retrieved for evaluation. In practices with electronic medical records, a random-number generator produced a list of 100 practice patients. In each case the chart abstractor reviewed eligibility sequentially until 30 eligible charts were identified for review.

Data collection tools. We used a theory-based evaluation framework to identify the dimensions of care that should be addressed and to help select the tools used for the evaluation.[5] The process involved a review of the literature and consultation with stakeholders and experts in the field to develop the theory underpinning the approach. As a result, we developed a conceptual framework that identified key areas to measure;[6] established program logic models for each practice model that provided a detailed visualization of the link between organizational attributes, activities and performance; and produced a mapping document to guide the tool selection.

Quantitative component. The quantitative data collection tools comprised 3 surveys and a chart abstraction form. The surveys were modified from the adult edition of the Primary Care Assessment Tool (PCAT), full or abridged version. The PCAT is an instrument developed to measure the quality of primary care services. The full version of the PCAT was validated in 2001.[11,12] We selected this tool because of the high degree of congruency between the dimensions it addresses and those set out in our conceptual framework and because the instrument allows the perceptions of patients and providers to be measured. To maintain the validity of the original tool, which was developed in the US, modifications were kept to a minimum and primarily reflected the differences in context between the US and Ontario settings. To minimize the burden on providers in group practices, a subset of questions from the provider survey addressing practice factors common to all of the providers in a given practice was moved to a practice survey.

The content of the PCAT was mapped to the dimensions of the conceptual framework, and where deficiencies were noted the tool was supplemented with questions from the National Physician Survey and other studies[9,13-15] or with questions developed by the investigators. Copies of the surveys are available from the authors upon request. Details of the scales and indicators used in this evaluation are shown in Table 1.

Practice survey. The practice survey was divided into 3 sections. The first focused on the description of the practice environment including the setting, hours of operation, availability of medical and social services in the surrounding area and accessibility for disabled persons. The second section contained questions that measured performance (see Table 1). The third section captured various practice attributes, including governance, team structure, extent of information technology adoption and economic information (e.g., sources of income, salaries and operating costs).

Provider survey. The provider survey was divided into 2 sections. The first section contained questions measuring the provider's perception of practice performance on several dimensions of health care service delivery (see Table 1). The second section captured provider demographic information, information on their work setting and socio-economic information.

Patient survey. The patient survey was divided into 2 sections. The first section was completed in the waiting room before the visit with the provider. This section captured patient sociodemographic and economic information and elicited the patient's experience concerning a broad range of dimensions of health care service delivery as shown in Table 1. The second section, completed after the appointment with the provider, took

less than 5 minutes to answer and captured visit-specific information, including waiting time, visit duration and measures of activities related to health promotion.

The survey was developed in English and translated to French through an extensive iterative translation process. The French version was validated against the English version on a sample of 120 bilingual individuals.[15] We made the tool available in French and English only and relied on the services of translators to reach patients who spoke neither language.

Chart audit. The chart audit forms captured 4 thematic areas: patient demographic information; visit activities, including referrals, prescriptions and orders; chart organization; and measures of performance of technical quality of care, including prevention, chronic disease management and acute disease management. We evaluated performance of technical quality of care by comparing the care provided with established guidelines for prevention, chronic disease management and acute disease management.

Qualitative component. We used the conceptual framework to define the topics and questions to be covered during qualitative data collection. At the case study sites at least 2 physicians and at least 1 nurse practitioner (if available) were interviewed. The interview guide for providers contained questions about the influence of organizational characteristics (e.g., remuneration scheme), processes (e.g., teamwork, inter-professional collaboration) and clinical routines on service delivery. The interview guide for patients focused on their experience with the practice associated with the dimensions of accessibility, continuity, coordination and comprehensiveness of care. The interviews with key informants focused on qualitative comparisons of the 4 models studied in relation to broad issues such as governance, accountability and performance measurement in primary care.

Quality control. All tools were piloted before the start of the study. A full description of the piloting process can be found in Appendix 1. Data entry verification was performed for all 4 tools, and the accuracy with which the results of the practice and provider survey were recorded was enhanced by double data entry. Chart audit validation was performed twice during the study. At each verification, chart abstractors were informed of their errors and received additional focused training then and throughout the study. Data were exported into SPSS and verified for internal consistency, missing information and outliers. Queried data were verified against the hard copy of the data collection tools. The validity of the qualitative findings was verified using naturalistic inquiries.[16] We also engaged in member-checking procedures to establish the credibility of our findings. Finally, the use of data triangulation techniques increased the construct validity of our measures for the performance domain (for both the quantitative and qualitative components). Additional details concerning the quality control processes are available in Appendix 1.

Study processes. This study involved a wide range of personnel from various backgrounds over a 3-year period and required significant organizational preparation. Details of the study team composition and study processes are available in Appendix 2.

Stakeholder advisory meeting. A stakeholder advisory committee comprised of 2 members from each model, Ministry of Health and Long-Term Care representatives, a community member and study team members met twice during the study. The committee's goals were for its members to serve as conduits between their representative group and the study team, to ensure transparency of the study process, to guide the evaluation plan and interpretation of results, and to participate in outcome dissemination.

Planned analyses. The study captures 2 types of data, 1 describing the practice structure and the other the practice performance (see Table 1). The study will use multi-level analyses to compare the performance of the models studied across the performance dimensions. It will also rely on the large number of structural attributes described for each practice to assess their impact on performance by evaluating their association with better performance. For example, we will evaluate whether a difference in first contact accessibility exists between models and then identify the components of the practice structure that are associated with better first contact accessibility across all models. In these analyses, provider information will be aggregated to the practice level, and patient level information (from surveys and chart abstraction) will be linked to the practice and provider data, allowing a hierarchical approach to data analysis accounting for intra-cluster correlations.[7] We captured measures of the quality of health service delivery as well as measures of the technical quality of care in the sample practices. Our analyses will also allow us to understand the relationship between the 2 within a practice.

Results

The study was successful in recruiting its intended number of practices (35) in all practice types except HSOs (32) (Table 2) and involved 8 practices in the qualitative evaluation. FFS–FHG practices were the most difficult ones to recruit (participation rate of 23%). We compared the profiles of the recruited family physicians with the profiles of all Ontario family physicians practising in these models to determine if there was selection bias related to practice refusal or provider

Table 3: Practice representativeness by model

Characteristic	FFS		FHN		CHC		HSO	
	Study *n = 58*	ICES *n = 9055*	Study *n = 80*	ICES *n = 590*	Study *n = 108*	ICES *n = 186*	Study *n = 42*	ICES *n = 165*
Provider demographic profile								
Male, %	52	66	60	65	42	51	74	73
Years since graduation, mean	22	26	23	22	19	19	29	28
Foreign trained, %	17.2	21.8	2.5	10.5	9.3	9.7	14.3	9.1
Rural, %	12	10	21	33	—	—	—	—
Provider work profile*								
Total visits, no.	5873	5389	4893	5329	—	—	—	—
Emergency visits, no.	265	199	180	331	—	—	—	—
Office visits, no.	5201	4651	4145	4279	—	—	—	—
Total payments (x 1000)	$192	$188	$140	$162	—	—	—	—

The sample sizes in the column headings represent the number of providers.

FFS = traditional fee-for-service and family health group, FHN = family health network, CHC = community health centre, HSO = health service organization, ICES = Institute for Clinical Evaluative Sciences.

* To create this profile, virtual patient rosters were built in the administrative databases in which patients were assigned to a practice if at least 50% of the billing associated with their care was submitted by a provider at that practice. These workload data were based on the period from Apr. 1, 2004, to Mar. 31, 2006.

Table 4: CHC patient representativeness

Sociodemographic factor	Study	CHC database
Age, yr (mean)	48.6	46.2
Female, %	71	50
Born in Canada, %	69	67
Duration in Canada, yr (mean)	18	13
Education, %		
No formal education	1	10
Primary or secondary	56	55
Post-secondary	43	36
Household income, %		
$0—$14 900	25	37
$15 000—$34 900	26	23
> $35 000	49	40
Single-person household, %	16	14
Insurance coverage, %		
Insured in Ontario (includes insured in Canada)	92	91
Uninsured	8	9

CHC = community health centre. Fourteen CHC practices participated in this evaluation. The table shows the sociodemographic factors of all patients in these 14 practices and of the patients participating in the study from these same practices.

Table 5: Extent of scale completion

Scale	% evaluable Pr	% evaluable Pt
Access		
First contact accessibility scales	100	94
First contact utilization scale	–	99
Patient-provider relationship		
Humanism scale	–	99
Trust scale	–	98
Cultural competency scales	99	89
Family centredness scales	100	96
Continuity		
Ongoing care scale	–	99
Service integration		
Coordination scale	100	–
Coordination: information system scale (PCAT)	100	–
Comprehensiveness		
Services offered (PCAT)	100	–
Community orientation scale: reach out	97	–
Community orientation scale: needs assessment	96	–
Community orientation scale: monitor	98	–
Community orientation scales	100	72
Provider satisfaction		
Provider satisfaction scale	100	–
Technical quality of clinical care delivery		
Health promotion	–	91
Overall average	99	93

Pr = scale included in the provider or practice survey, Pt = scale completed by the patient, PCAT = Primary Care Assessment Tool.

self-selection. We relied on the information contained in the physician workforce database and in the Ontario Health Insurance Plan (OHIP) billing database housed at ICES. The former allowed evaluation of provider demographic profiles, and the latter provided billing parameters that allowed us to compare the FFS–FHG and FHN practices only (these models rely on Ministry of Health and Long-Term Care billing for their remuneration). These comparisons showed that our sample is broadly representative for all characteristics measured in these databases (Table 3).

We compared the sociodemographic information of the CHC patients participating in the study with that of all CHC patients listed in the CHC practice electronic patient registration database to evaluate whether there was systematic bias in the selection of respondents from the CHCs (Table 4). CHC is the model most likely to serve individuals who are housebound or have language barriers and therefore less likely to have been reached in this study than patients from the other practice types. As anticipated, the waiting room sample was older and more likely to be female than the overall practice population, reflecting the profile of those who make more use of primary care services. The study sampling was not successful in reaching individuals without a formal education and those with lower income.

Survey questionnaires were not modified after the start of the study. All practices and all but 2 consenting providers completed the survey. The overall patient participation rate was 82%, with most scales adequately completed for evaluation (Table 5).

Discussion

We measured performance across a large number of primary care attributes to obtain a comprehensive picture of status of family care in Ontario. We evaluated dimensions of health service delivery and technical quality of care in the same practices. The study was complex and care was taken to ensure the quality of the data collected and to minimize disruption to the practices. At the study onset, much work was invested in ensuring that appropriate evaluation tools were used. Throughout the study, we focused on enhancing practice and patient recruitment, establishing dependable processes for data collection, verifying data quality and training and supporting personnel.

The study was successful in collecting data from 137 primary care practices for a multi-dimensional evaluation. The limitations of this mixed-methods study stem largely from the problems inherent in cross-sectional and survey-based studies. These include participant selection bias and the inability to infer causation from observed associations. Other study-specific factors are discussed below.

Sample selection. Sample selection was limited by our ability to identify all practices within a model, the geographic boundaries we established for data collection and the fact that patient recruitment was limited to those attending the practice. There was no accessible central source of reliable practice lists within each model, except for CHCs. In addition, late in 2004 the Ontario Ministry of Health and Long-Term Care instituted a new model of care, the FHG, to which FFS practices could transition. We initially excluded FHG practices, but FFS practices converted to this new model quickly; by early 2006 most FFS practices had become FHGs and it became evident that the great majority would transition by the year end. As a result, 3 months before recruitment was terminated, a decision was made to include the FFS practices that had transitioned to FHGs. Although a concerted effort was made to return to those practices initially deemed ineligible because they had converted into an FHG, not all attempts were successful, so we cannot ignore this potential source of bias toward late adopters within this subset.

The geographic boundaries set by the study resulted in the exclusion of the most northern territories of the province. These areas serve a more marginalized population living under very different conditions and for whom the experience of primary care services is not reflected by the study sample. Our study's findings cannot be extrapolated to that group.

Finally, we chose to administer the patient survey to those patients visiting the practice on a given day. This face-to-face approach is expected to have enhanced our response rate (compared with what might have been expected with a telephone or mailed questionnaire approach) but resulted in an overrepresentation of those more likely to frequent the practice. Therefore, the sample does not represent the general practice population, nor did it reach housebound patients. Rather it is weighted, perhaps appropriately so, by the frequency of visits.

In contrast, the chart-based assessment of the technical quality of care was based on a random selection of records so that the results could be generalizable to the practice level. An alternative strategy would have been to review the charts associated with the patients surveyed. Although that approach would have allowed the relationship between the quality of health service delivery and technical quality of care to be assessed at the individual patient level, the estimates of care level would have been biased toward those attending the practice more frequently.

Data. Although the original PCAT tool had been validated,[12] for some scales we relied on the nonvalidated abridged version of a validated scale. We made the tool available in 2 languages only (French and English) and used the services of translators to reach patients who spoke neither of these languages. Although we felt it was essential to capture the essence of the experience of patients from linguistic minority groups, the use of an intermediary allows for biases or inconsistencies to be introduced during the translation process.

Ideally, the selection of practices for the case study would have been informed by the results of the quantitative surveys concerning the quality-of-care indicators. This would have allowed us to select negative or deviant cases within each model for in-depth analyses. However, because of time constraints, sites were invited to participate in both components (quantitative and qualitative) of the onset of the study.

Participation. This study was conducted at a time when Ontario primary care practices were saturated with government-sponsored studies, which likely contributed to the suboptimal participation rate. The practice response rate was best in models from which we obtained support from their central organizational group (CHC and HSO). Despite lower participation rates in FFS–FHG and FHN practices, comparative data suggest that the study population was adequately representative. All but 1 scale had completion rates of 94% or higher.

We compared the study patient population with the general practice population in CHCs and found that CHC participants were older, more likely to be female, had completed a higher level of education and had a higher income than the general CHC population. In Canada older people, women and people with higher socio-economic status are more likely to visit their family physician, and thus these differences between the CHC patients surveyed and those served in CHCs may be related to our waiting room sampling approach rather than participation bias.[17]

Conclusions

This is the first comprehensive pan-Ontario evaluation of models of primary care. The breadth of data collected will allow an in-depth description of the practices belonging to each model type. An evaluation of the practice factors (organizational features and practice attributes) associated with better performing practices should help inform policy-makers about optimal features in primary care practices and shoud help inform practice managers about how best to structure their practices to serve their disadvantaged patients. This article may also be useful to researchers interested in investigating issues related to quality of care and organizational performance in primary care.

Contributors: Simone Dahrouge participated in finalizing the study methodology, managed the quantitative component and was the principal writer of the manuscript. William Hogg conceived the project, oversaw the data collection and analysis and participated in all phases of the writing. Grant Russell helped implement the study, worked on finalizing the methodology and contributed to the writing and editing of the manuscript. Robert Geneau described the qualitative methods used in the study and reviewed all manuscript drafts. Elizabeth Kristjansson participated in editing and reviewing manuscript drafts. Laura Muldoon conceived the study and oversaw its implementation and participated in the writing of the manuscript. Sharon Johnston helped guide the analysis and participated in the writing. All of the authors approved the final version of the manuscript.

References

1. Rosenthal MB. Beyond pay for performance — emerging models of provider-payment reform. *N Engl J Med*. 2008;359(12):1197–1200.
2. Buetow S. Pay-for-performance in New Zealand primary health care. *J Health Organ Manag*. 2008;22(1):36–47.
3. Campbell S, Reeves D, Kontopantelis E, Middleton E, Sibbald B, Roland M. Quality of primary care in England with the introduction of pay for performance. *N Engl J Med*. 2007;357(2):181–190.
4. Pink GH, Brown AD, Studer ML, Reiter KL, Leatt P. Pay-for-performance in publicly financed healthcare: some international experience and considerations for Canada. *Healthc Pap*. 2006;6(4):8–26.
5. Rowan MS, Hogg W, Labrecque L, Kristjansson EA, Dahrouge S. A theory-based evaluation framework for primary care: setting the stage to evaluate the Comparison of Models of Primary Health Care in Ontario (COMP-PC) project. *Can J Program Eval*. 2008;23(1):113–140.
6. Hogg W, Rowan M, Russell G, Geneau R, Muldoon L. Framework for primary care organizations: the importance of a structural domain. *Int J Qual Health Care*. 2007;20(5):308–313.
7. Baskerville NB, Hogg W, Lemelin J. The effect of cluster randomization on sample size in prevention research. *J Fam Pract*. 2001;50(3):W241–W246.
8. Ontario Medical Association*Family health groups (FHGs): Frequently asked questions*. Toronto: The Association: 2007.
9. College of Family Physicians of Canada*Updated data release of the 2001 National Family Physician Workforce Survey*. 2002 [accessed 7 Mar 2007].
10. Patton MQ. *Qualitative research and evaluation methods*. Thousand Oaks (CA): Sage Publications, Inc: 2002.
11. Starfield B. *Adult primary care assessment tool—expanded version*. Boston: Primary Care Policy Center, Johns Hopkins University School of Hygiene and Public Health: 1998.
12. Shi L, Starfield B, Xu J. Validating the adult primary care assessment tool. *J Fam Pract*. 2001;50(2):E1.
13. Hauck FR, Zyzanski SJ, Alemagno SA, Medalie JH. Patient perceptions of humanism in physicians: effects on positive health behaviors. *Fam Med*. 1990;22(6):447–452.
14. Hall MA, Zheng B, Dugan E, Camacho F, Kidd KE, Mishra A, et al. Measuring patients' trust in their primary care providers. *Med Care Res Rev*. 2002;59(3):293–318.
15. Haggerty J. Centre de recherche du Centre hospitalier de l'Université de Montréal (CHUM)*Accessibility and continuity of primary care in Quebec. Annex 2: Primary Care Assessment Questionnaire*. 2004 [accessed 14 Aug 2009].
16. Lincoln YS, Guba EG. *Naturalistic inquiry*. Newbury Park (CA): Sage Publications, Inc: 1985.
17. Seale C. *The quality of qualitative research*. London: Sage Publications, Inc: 1999.
18. Nabalamba A, Millar WJ. Going to the doctor. *Health Rep*. 2007;18(1):23–35.

Citation: Dahrouge S, Hogg W, Russell G, Geneau R, Kristjansson E, Muldoon L, Johnston S. The Comparison of Models of Primary Care in Ontario (COMP-PC) study: methodology of a multifaceted cross-sectional practice-based study. *Open Med* 2009;3(3):149-164

Published: 1 September 2009

Appendix 1: Quality control processes

Tool piloting

Drafts of the surveys were reviewed and piloted iteratively by members of the study team until general consensus was reached about tool readiness. External piloting of the surveys and chart audit took place between July and October 2005 and involved 6 practices in the Ottawa area: 2 FFS practices, 2 FHGs, 1 FHN and 1 CHC. The practices were recruited sequentially, allowing for adjustments to be made to the study tools in response to the results from the previous site piloted. At each site 1 practice survey was completed, 30 patients were surveyed and 30 charts were reviewed. In total, 18 consenting physicians completed the provider survey.

The piloting process identified a number of issues that allowed us to refine the chart audit, to clarify some survey questions and to streamline the process. Most of the changes were to the questions that had been added to the instrument by the study team. We modified a validated scale only when absolutely necessary. Smaller numbers of concerns were raised at successive iterations of the tools; the iterative process was stopped when no new issues were identified. Two of the pilot sites (1 CHC and 1 FHN) were also visited by the qualitative researchers to validate the provider and patient interview guides.

Validation of survey data entry

Survey administrators administered the patient surveys and entered the data into a web-based program. In certain sites, they also recruited patients for the qualitative component. Chart abstractors were responsible for distributing the provider and practice surveys and performing the chart audit and its web-based data entry. Shortly after we initiated data collection, a research associate performed duplicate data entry for patient surveys and chart audits for 8 sites to estimate data-entry error rates for all chart abstractors and survey administrators. Error rates were 1.3%, 1.4%, 4.4% and 0.33% for patient, provider and practice surveys and chart audits, respectively. A substantial proportion of the discrepancies observed in the survey data related to the assignment of "0" by the survey administrators in numeric fields where no data had been recorded on the form. Detailed feedback and instructions were provided to all personnel involved in data entry. To ensure that high-quality data were collected, at the end of the study the data from all provider and practice surveys were re-entered by another research associate (double data entry) and errors were corrected. A final random verification of the data entry for the patient surveys found that the overall error rate had dropped to 0.5%.

Chart audit validation

Seven chart abstractors performed chart abstractions at 137 practices. Because this process is rather complex and prone to human error, a review process was set up to ensure the quality and consistency of data extraction. The abstractors were informed that validation would take place throughout the study and were required to maintain for that purpose a list of the charts they had reviewed at each practice. Validation involved duplication of the entire data extraction for 8 charts. We defined levels of error and took action according to the extent to which these errors were observed. An error that led to the failure to recognize eligibility for more than one manoeuvre (e.g., age miscalculated, chronic disease not recognized) was considered most significant (level 1). A level 2 error was defined as a missing visit record or incorrect attribution of eligibility that led to a single manoeuvre or sub-question being missed. All other coding errors were considered minor. The presence of at least one level 1 error or two level 2 errors was considered a significant problem and led to the validation of all the remaining charts that had been reviewed at that practice. Errors encountered were corrected on the data collection form. The chart abstractors were informed of their errors, and this opportunity was used for further general training. The abstractors also received ongoing training and support throughout the chart abstraction process.

The first round of validation was performed very shortly after the start of the study and involved the first 2 sites completed by each of the 7 chart abstractors. At that time, level 2 errors (at least 2) were identified in the charts of 2 abstractors. The second round of validation occurred during the winter of 2006. At that time, 6 abstractors were active; again 2 sites were randomly selected for each abstractor and the same procedure was followed. No significant problems were encountered.

Qualitative component

We used several procedures to ensure the validity of the qualitative findings. We adopted the criteria defined by Lincoln and Guba for naturalistic inquiries.[16] First, we offered a detailed description of each site and of each primary health care model. Providing a thick description is one of the key characteristics leading to the transferability of qualitative results to other contexts.[17] Second, we asked external peer reviewers to critique and challenge our research design, analytic strategy and interpretation of findings to increase the level of dependability of our study. We also engaged in member-checking procedures to establish the credibility of our findings by sending our case study reports to members of the 8 participating sites. The key informants were also asked to review a draft copy of the final report. Finally, the use of data triangulation techniques increased the construct validity of our measures for the performance domain (for both the quantitative and qualitative components). We blended qualitative and quantitative approaches to answer the research questions and reach consistent and valid conclusions. This approach also allowed us to verify the validity of the constructs addressed in the quantitative survey.

Appendix 2: Study processes

Two goals directed our study process: data quality and minimal practice disruption.

Data quality

Personnel training. The survey administrators received a half-day of training at a central location and were then paired with an experienced survey administrator for 2–3 days of fieldwork at 1 practice. The chart abstractors received 2 days of training at a central location and then carried out 1 day of fieldwork with an experienced chart abstractor. Each group was guided by a detailed instruction manual. Data collectors were also provided with a toll-free telephone number for the project team so that they could call if they had any questions or needed to report problems encountered in the field. Instruction manuals were revised periodically to reflect new information and were re-disseminated to the data collectors.

Data collection processes. At the first visit to the practice, the chart abstractor met with the office manager (and, when possible, the participating providers) to distribute and review the content of the practice and provider surveys and to offer assistance in interpreting the questions. If a survey was not completed during the data collection period in the practice, it was left with the respondent together with a cover letter and a self-addressed, postage-paid return envelope. A research assistant telephoned non-responders after 2 weeks, and, when required, this was followed by the mailing of a second (and, if necessary, third) copy of the survey package. The site received financial compensation only when all surveys had been returned and data collection completed. To increase the rate of completion of patient surveys, the survey administrators were available to answer patient questions and ensure that patients leaving the office completed the post-visit survey.

Chart abstractors were required to familiarize themselves with the charting system of each practice and inquire about all potential sources of clinical information, including electronic and paper-based medical records, registries of influenza vaccinations, medication lists and laboratory results.

All patients and providers participating in semi-structured interviews had first completed the survey. Our early access to quantitative data allowed us to customize to some extent the interview guide for each respondent. The providers' interviews focused on interrelationships between the organizational structures and processes, the practice context and the various dimensions of quality of care. The interviews with patients explored their experience receiving health care whereas those with key informants emphasized macro level issues such as governance and accountability. The average interview duration was 90 minutes.

Data entry. We used a web-based clinical data management tool (TrialStat Corporation, Ottawa, Ont.) to store all quantitative data. The customized electronic data capture forms contained rules for data entry validation (ranges, missing information and internal consistency [i.e., congruence between the data entered in related fields]) so as to minimize data entry errors. The survey administrators and chart abstractors entered the results of the surveys and chart audits, respectively, into the system from remote locations. Variable fields, labels and data were then exported directly into SPSS (SPSS Inc., Chicago, Ill.) for analysis. Data entry validation (targeting all chart abstractors and survey administrators) was performed by a research associate. All qualitative interviews were transcribed verbatim using Microsoft Word and were then validated and imported into N6 (QSR International, Doncaster, Victoria, Australia) for analysis.

Minimum disruption to the practice

To minimize disruption to the practice and to ensure a seamless effort, the data collectors received relevant information for each site, including the type of facility, contact information, participating providers and the best day and time to call. The chart abstractor and survey administrator coordinated the logistics of their data collection efforts before contacting the site. One team member was assigned to be the point of contact with the practice, and the 2 team members maintained contact with each other throughout the data collection period. Survey administration and chart abstraction required an average of 31 and 20 hours per practice, respectively.