

Measurement of Primary Healthcare Attributes from the Patient Perspective

Mesure des caractéristiques des soins de santé primaires du point de vue du patient

JEANNIE L. HAGGERTY, PHD
*Department of Family Medicine, McGill University
Montreal, QC*

THIS SPECIAL ISSUE OF *HEALTHCARE POLICY/POLITIQUES DE SANTÉ* PRESENTS A series of papers reporting on the concurrent validation of instruments that assess primary healthcare (PHC) delivery from the patient's perspective. The study was funded in 2004 by the Canadian Institutes of Health Research (CIHR) at the height of the Primary Health Care Transition Fund, an \$8-million investment by Health Canada to catalyze a renewal of the PHC system in Canada. In planning the evaluation for initiatives, it became evident that program evaluators and researchers had little guidance for selecting among available instruments. For example, although various instruments purported to measure accessibility, it was not obvious that all measured the same underlying construct. Nor was it clear how to compare results collected with different instruments in different jurisdictions or at different times.

This research program consists of three studies. The first is a consensus consultation of PHC experts across Canada to formulate operational definitions of attributes to be evaluated (Haggerty et al. 2007). Second, we mapped the operational definitions to validated instruments (available at www.programmeprecise.ca/en/publications). Third, we administered six instruments back-to-back in Nova Scotia and Quebec to adults with a regular source of care to examine

and compare how well the instruments measure essential attributes of primary healthcare.

We selected six instruments in the public domain that assess usual care rather than a single visit, that are generic (not limited to a specific patient group or dimension of care) and that had been most proposed or used in Canada:

1. Primary Care Assessment Survey (PCAS) (Safran et al. 1998);
2. Primary Care Assessment Tool – Short Form, adult (PCAT-S) (Shi et al. 2001);
3. Components of Primary Care Index (CPCI) (Flocke 1997);
4. First version of the European general practice evaluation instrument (EUROPEP-I) (Grol et al. 2000);
5. Interpersonal Processes of Care – 18-item version (IPC-II) survey (Stewart et al. 2007);
6. Veterans Affairs National Outpatient Customer Satisfaction Survey (VANOCSS) (Borowsky et al. 2002).

Overview of the Findings

The first paper in this special issue reports how primary healthcare experts prioritized attributes for different primary healthcare models (Lévesque, Haggerty et al. 2011). Some attributes were identified as essential to any primary healthcare model; others assumed greater or lesser importance if the model was professional or community-oriented.

The methods article (Haggerty, Burge, Beaulieu et al. 2011) describes the sampling, recruitment and descriptive results of the concurrent instrument administration. Most instrument subscales discriminate among poor, average and excellent overall experience, but the most discriminating attributes are interpersonal communication and respectfulness. Inclusion of these attributes underlines their importance to patients and the crucial value of good measures.

Our sample was balanced by French/English language, high/low education, urban/rural location and overall healthcare experience, so we had sufficient statistical power to assess whether instruments perform differentially (are biased) by these categories. Because we administered French instruments in one province and English in another, we used advanced statistics to determine whether observed differences were due to language or the health system. The paper on differential item functioning (Haggerty, Bouharaoui et al. 2011) presents an overview of this technique and demonstrates that differential functioning was prevalent between French- and English-language instruments. Using only unbiased questions, we found that primary healthcare experience was consistently better in Nova Scotia than in Quebec. Also, with unbiased questions, we reversed an initial finding and found instead that first-contact accessibility was significantly worse for rural than urban respondents.

We learned qualitatively through 13 discussion groups (Haggerty, Beaulieu, Lawson et al. 2011) that patients overwhelmingly prefer Likert response scales with labels adapted to the context of the question. They do not like agree–disagree options to elicit frequency. They prefer long and clear formats to those that are short and crowded. Patients want to be good respondents and admit to guessing when they cannot evaluate directly or don't understand the

question. They were disappointed that the instruments did not allow them to report on problems with the interface between different providers in the system.

The core question in our study, however, was how well different instruments measure the attributes essential to primary healthcare: accessibility, interpersonal communication, comprehensiveness, relational continuity, management continuity and respectfulness. All the study instruments have passed standards of reliability and validity, but we went further by comparing subscale values using a common metric, conducting factor analysis of items from different instruments, and then examining how well individual items measured the common construct that emerged across instruments. These findings are outlined and explained in the analytic overview paper (Santor et al. 2011). Pooling items allowed us to discern different dimensions within an attribute, some within the same instrument subscale.

Accessibility has been identified as a weakness in Canadian health systems in sequential international surveys (Schoen et al. 2004, 2007). It is therefore good to know that two subscales perform well (Haggerty, Lévesque et al. 2011). The PCAT-S First-Contact Access, despite some measurement problems, is the best and only measure of patients' confidence in being seen rapidly. The PCAS Organizational Accessibility has good metric properties and measures accommodation rather than rapid access.

Interpersonal communication – the heart of patient-centred care and foundational to establishing a therapeutic alliance – can suffer in team-based and shared care (Rodriguez et al. 2007; Safran 2003), so monitoring is critical. We discerned sub-dimensions of eliciting, explaining and shared decision-making that are captured in the PCAS Communication and the EUROPEP-I Clinical Behaviour subscales, with the former showing better measurement properties (Beaulieu et al. 2011). Three subscales in the IPC-II measure these dimensions specifically, but the response options or scoring could be adjusted to permit better discriminability.

Relational continuity is valued highly in family medicine as having therapeutic potential in itself, and is a potential victim of reforms towards team-based care. The assumption that continuity is achieved through concentrating care in a single physician is reflected in the measures we studied (Burge et al. 2011), though scores and percentage of visits with doctor do not correlate highly. This attribute is inferred from accumulated and comprehensive knowledge of the patient, captured in two subscales with similar content, the CPCI Accumulated Knowledge and the PCAS Contextual Knowledge, the latter with better metric properties. These may also measure an aspect of whole-person care.

Comprehensiveness is one of the most invoked qualifiers of good primary healthcare, but the lack of definitional clarity complicates measurement (Haggerty, Beaulieu, Pineault et al. 2011). Comprehensiveness as whole-person care is missing from these measures and needs development. The CPCI Comprehensive Care measures the patient's confidence in the physician's capacity to care for a range of health problems but may not reflect the actual range of services. The PCAT-S Comprehensive Services Available elicits the range of services, but measurement is limited by patient knowledge, needs or both. Range of services is probably best assessed by providers.

Management continuity is the experience of care coordination. Only some provider efforts to link and coordinate care are visible to patients, a fact that may explain why patient assessments of their primary care physician's coordination actions with the PCAT-S Coordination and the PCAS Integration are predominantly positive, even when problems are reported (Haggerty, Burge, Pineault et al. 2011). Participants in discussion groups wanted to report on their care experience across the entire system (Haggerty, Beaulieu et al. 2011), and the VANOCSS Overall Coordination subscale is the best tool to capture this experience.

Respectfulness – attention to dignity, interpersonal treatment and adequate privacy – is a new dimension, important to patients and their confidence in the health system. Respectfulness is the principal way patients experience responsiveness, defined as a fundamental dimension of health system performance by the World Health Organization and in which Canada showed room for improvement (WHO 2000). Though the paper by Lévesque, Pineault and colleagues (2011) is more exploratory than others in this series, they found respectfulness items in various instruments: addressed explicitly in IPC-II subscales but also by the PCAS Interpersonal Treatment. The measures apply to organizational processes as well as the clinical encounter.

Across attributes, patient assessments of care typically have a skewed distribution, with the vast majority of patients endorsing the more positive response options. This deviation from a normal distribution compromises the robustness of some psychometric measures, but most problematically, it reduces the capacity to discriminate between different levels of positive experience and, hence, to detect improvements. Patients are reluctant to evaluate providers negatively unless they know that the provider is responsible for the negative experience (Collins and O' Cathain 2003), suggesting that negative assessments are true negatives, whereas some positive assessments will be false positives. Indeed, item response analysis showed enhanced discriminability and information yield in the negative zone of assessments. This finding has implications for reporting of measures: rather than averaging item values, it may be more meaningful and informative to dichotomize assessments and to report the percentage with negative evaluations.

Patient assessments may be “contaminated” by the regard – usually positive – for the physician, and we may need new approaches to isolate experience of the attributes themselves. Instruments such as the VANOCSS, inspired by the Picker Institute suite of tools, elicit patient reports of incidents and preclude patients' judging their provider. This approach may help provide more information about positive experience. However, because it does not lend itself to classic psychometric analysis, we may need a new approach to assessing these instruments' validity and reliability.

Discussion

The results reported in the various papers in this special issue will provide guidance in the selection of instruments to be used in evaluation of primary healthcare and in the refinement and development of other instruments. Despite the end of the Primary Health Care Transition Fund in 2006, the results are still relevant because PHC renewal continues.

In fact, with the withdrawal of special funding, it is more important than ever to base resource allocations on sound evaluations and evidence of impact. These results are relevant not only for Canada, but also internationally – for example, to evaluate the patient-centred medical home (PCMH) in the United States and to inform the surveys of patient experience that are part of the pay-for-performance Quality and Outcomes Framework in the United Kingdom.

Ultimately, however, the enduring relevance of this study comes from the patient's being the *raison d'être* of the healthcare enterprise. Patient-centredness is espoused as a core value, but without mechanisms to continually remind us of how patients experience care, care delivery too easily becomes provider-centred. Patient surveys are critical to continuous quality improvement and the greater health system accountability promised to Canadians in the 2003 Health Accord.

In these papers, we refer unabashedly to “patients” rather than “consumers” or “clients,” even though use of surveys is most often associated with the market concerns that these latter terms evoke. The etymological root for “patient” is the Latin *pati*: “to suffer.” To “endure in suffering” constitutes the moral virtue of patience, one that is too often required in the healthcare delivery process as well as on the journey from illness to recovery. It is our sincere hope that this study will help evaluators institute mechanisms to listen effectively to patients' voices and adapt to their concerns.

Limitations of our study

This unique study gave us the rare privilege of comparing measures directly. However, there are some important limitations and things we might do differently were we to repeat the study.

First, few of the findings are generalizable beyond the six instruments included in our study. Resource limitations and the burden of response imposed the sometimes painful necessity of selecting instruments we believed to be most relevant and those that measured the usual care experience. Consequently, we selected the EUROPEP-I over the Consumer Assessment of Health Plan Survey (CAHPS, www.cahps.ahrq.gov), which is widely used in the United States, because the former fits our tax-based healthcare system. Likewise, we did not include any visit-based measures that, in retrospect, may be the most precise way to measure accessibility, interpersonal communication or respectfulness.

Second, we did not measure indicators of intermediate outcomes. Thus, we cannot estimate the relative association between, say, accessibility and unmet needs for care or between management continuity and gaps in information transfer.

Third, while most reliability and validity parameters of the instruments are similar in the Canadian context to the context in which they were developed, there are still unacceptably high levels of missing values in some items and in some instruments. Some of these are due to offering patients “not applicable” response options, which were appreciated by respondents but are essentially non-informative. These missing values limited our statistical power despite our having a robust sample size of 645 respondents. Evaluators should be alert to this pitfall in selecting instruments, as it could compromise the information yield of administered instruments.

Recommendations for researchers, program evaluators and policy makers

We did not aim to evaluate the instruments per se, but because our opinions are frequently solicited, I venture some observations.

The PCAS seems to have the best measurement properties overall: the response scales behave as expected, and the formatting and readability enhance ease of response. However, the questions pertain to the “regular personal doctor” and are not specific to primary healthcare – indeed, the content could well apply to specialist care. In contrast, the PCAT-S and the CPCI were specifically designed for and apply to North American primary healthcare; both have measurement problems that could be ameliorated easily by formatting to ease response and by adjusting the response options to better fit the questions. The EUROPEP-I, designed for European general practice, is easy to answer and its broad use allows for benchmarking; but it has only two subscales, which capture generic experience but present difficulties for monitoring specific attributes. The IPC-II measures dimensions that are important to patients and can be focused on primary healthcare, but it was not designed specifically for this; its response scale is not fully exploited, and the scoring may mask negative experiences.

Researchers need to develop measures for such attributes as whole-person care, shared decision-making, team relational continuity, information management and cultural sensitivity. Studies need to link patient assessments of attributes to intermediate and health outcomes. Finally, some fundamental methodological development is needed to assess the metric performance of instruments that elicit reports of critical incidents as a way of inferring quality of care.

We are confident in endorsing the instruments in our study for the Canadian context, despite their having been developed elsewhere. Program evaluators should, however, be cautious about subscales having high (over 4%) missing values, as these will compromise the capacity to detect intended effects. Our study identified sub-dimensions within the attributes being measured; these would be a consideration in selecting instruments. Data from providers is needed for such attributes as comprehensive services, integration and quality of care. Finally, we caution against using our results to select individual items for use in an isolated manner; we affirm the integrity of subscales as constructed by the developers.

Finally, policy makers may be assured that we have adequate measurement instruments to consistently and routinely measure patient perceptions of such attributes as accessibility, interpersonal communication, relational continuity and respectfulness. These attributes are important to Canadians and are predictive of their overall experience of care. Even if different instruments are used across time or across jurisdictions, our study demonstrates how a common metric can be established to compare results.

ACKNOWLEDGEMENTS

This research was funded by the Canadian Institutes of Health Research. This special supplement was sponsored by: a dissemination grant from the Canadian Institutes for Health Research; funds allocated to Jeannie Haggerty’s Canada Research Chair by the Université

de Sherbrooke and the Centre de recherche de l'Hôpital Charles LeMoyné; the Réseau de recherche en santé de populations du Québec and the Canadian Institute for Health Information. The production of an extended report was made possible by contributions from the following authors' institutions: Centre de recherche du Centre hospitalier de l'Université de Montréal; the McGill Research Chair in Family & Community Medicine at St. Mary's Hospital; Dalhousie University's Department of Family Medicine; University of Ottawa's School of Psychology and the Département de santé publique de l'Agence de santé et services sociaux de Montréal-Centre.

Correspondence may be directed to: Jeannie L. Haggerty, Associate Professor, Department of Family Medicine, McGill University, St. Mary's Research Centre, Hayes Pavilion – Suite 3734, 3830 Lacombe Ave., Montreal QC H3T 1M5; tel.: 514-345-3511 ext. 6332; fax: 514-734-2652; e-mail: jeannie.haggerty@mcgill.ca.

REFERENCES

- Beaulieu, M.-D., J.L. Haggerty, C. Beaulieu, F. Bouharaoui, J.-F. Lévesque, R. Pineault et al. 2011. "Interpersonal Communication from the Patient Perspective: Comparison of Primary Healthcare Evaluation Instruments." *Healthcare Policy 7* (Special Issue): 108–23.
- Borowsky, S.J., D.B. Nelson, J.C. Fortney, A.N. Hedeon, J.L. Bradley and M.K. Chapko. 2002. "VA Community-Based Outpatient Clinics: Performance Measures Based on Patient Perceptions of Care." *Medical Care 40*(7): 578–86.
- Burge, F., J.L. Haggerty, R. Pineault, M.-D. Beaulieu, J.-F. Lévesque, C. Beaulieu et al. 2011. "Relational Continuity from the Patient Perspective: Comparison of Primary Healthcare Evaluation Instruments." *Healthcare Policy 7* (Special Issue): 124–38.
- Collins, K. and A. O'Cathain. 2003. "The Continuum of Patient Satisfaction – From Satisfied to Very Satisfied." *Social Science and Medicine 57*(12): 2465–70.
- Flocke, S. 1997. "Measuring Attributes of Primary Care: Development of a New Instrument." *Journal of Family Practice 45*(1): 64–74.
- Grol, R., M. Wensing and Task Force on Patient Evaluations of General Practice. 2000. "Patients Evaluate General/Family Practice: The EUROPEP Instrument." Nijmegen, Netherlands: Centre for Quality of Care Research, Raboud University.
- Haggerty, J.L., M.-D. Beaulieu, R. Pineault, F. Burge, J.-F. Lévesque, D. A. Santor et al. 2011. "Comprehensiveness of Care from the Patient Perspective: Comparison of Primary Healthcare Evaluation Instruments." *Healthcare Policy 7* (Special Issue): 154–66.
- Haggerty, J.L., C. Beaulieu, B. Lawson, D.A. Santor, M. Fournier and F. Burge. 2011. "What Patients Tell Us about Primary Healthcare Evaluation Instruments: Response Formats, Bad Questions and Missing Pieces." *Healthcare Policy 7* (Special Issue): 66–78.
- Haggerty, J.L., F. Bouharaoui and D.A. Santor. 2011. "Differential Item Functioning in Primary Healthcare Evaluation Instruments by French/English Version, Educational Level and Urban/Rural Location." *Healthcare Policy 7* (Special Issue): 47–65.
- Haggerty, J.L., F. Burge, M.-D. Beaulieu, R. Pineault, C. Beaulieu et al. 2011. "Validation of Instruments to Evaluate Primary Healthcare from the Patient Perspective: Overview of the Method." *Healthcare Policy 7* (Special Issue): 31–46.

- Haggerty, J.L., F. Burge, J.-F. Lévesque, D. Gass, R. Pineault, M. Beaulieu et al. 2007. "Operational Definitions of Attributes of Primary Health Care: Consensus among Canadian Experts." *Annals of Family Medicine* 5(4): 336–44.
- Haggerty, J.L., F. Burge, R. Pineault, M.-D. Beaulieu, F. Bouharaoui, C. Beaulieu et al. 2011. "Management Continuity from the Patient Perspective: Comparison of Primary Healthcare Evaluation Instruments." *Healthcare Policy* 7 (Special Issue): 139–53.
- Haggerty, J.L., J.-F. Lévesque, D.A. Santor, F. Burge, C. Beaulieu, F. Bouharaoui et al. 2011. "Accessibility from the Patient Perspective: Comparison of Primary Healthcare Evaluation Instruments." *Healthcare Policy* 7 (Special Issue): 94–107.
- Lévesque, J.-F., J.L. Haggerty, F. Burge, M.-D. Beaulieu, D. Gass, R. Pineault et al. 2011. "Canadian Experts' Views on the Importance of Attributes within Professional and Community-Oriented Primary Healthcare Models." *Healthcare Policy* 7 (Special Issue): 21–30.
- Lévesque, J.-F., R. Pineault, J.L. Haggerty, F. Burge, M.-D. Beaulieu, D. Gass et al. 2011. "Respectfulness from the Patient Perspective: Comparison of Primary Healthcare Evaluation Instruments." *Healthcare Policy* 7 (Special Issue): 167–79.
- Rodriguez, H.P., W.H. Rogers, R.E. Marshall and D.G. Safran. 2007. "Multidisciplinary Primary Care Teams. Effects on the Quality of Clinician–Patient Interactions and Organizational Features of Care." *Medical Care* 45(1): 19–27.
- Safran, D.G. 2003. "Defining the Future of Primary Care: What Can We Learn from Patients?" *Annals of Internal Medicine* 138(3): 248–55.
- Safran, D.G., J. Kosinski, A.R. Tarlov, W.H. Rogers, D.A. Taira, N. Lieberman and J.E. Ware. 1998. "The Primary Care Assessment Survey: Tests of Data Quality and Measurement Performance." *Medical Care* 36(5): 728–39.
- Santor, D.A., J.L. Haggerty, J.-F. Lévesque, F. Burge, M.-D. Beaulieu, D. Gass et al. 2011. "An Overview of Confirmatory Factor Analysis and Item Response Analysis Applied to Instruments to Evaluate Primary Healthcare." *Healthcare Policy* 7 (Special Issue): 79–92.
- Schoen, C., R. Osborn, M. Doty, M. Bishop, J. Peugh and N. Murukutla. 2007. "Toward Higher-Performance Health Systems: Adults' Health Care Experiences in Seven Countries." *Health Affairs* 26(6): 717–34.
- Schoen, C., R. Osborn, P.T. Huynh, M. Doty, K. Davis, K. Zapert et al. 2004. "Primary Care and Health System Performance: Adults' Experiences in Five Countries." *Health Affairs* (Millwood), Suppl. Web Exclusives: W4–503.
- Shi, L., B. Starfield and J. Xu. 2001. "Validating the Adult Primary Care Assessment Tool." *Journal of Family Practice* 50(2): n161w–n171w.
- Stewart, A.L., A.M. Nápoles-Springer, S.E. Gregorich and J. Santoyo-Olsson. 2007. "Interpersonal Processes of Care Survey: Patient-Reported Measures for Diverse Groups." *Health Services Research* 42(3 Pt. 1): 1235–56.
- World Health Organization (WHO). 2000. *The World Health Report 2000 – Health Systems: Improving Performance*. Geneva: Author.