Measuring Consumer Experiences With Primary Care

Charlyn E. Cassady, PhD*; Barbara Starfield, MD, MPH**; Margarita P. Hurtado, PhD, MA, MHS§; Ronald A. Berk, PhD‡; Joy P. Nanda, MS, MHS*; and Lori A. Friedenberg, BA*

Abstract. Objectives. To assess the adequacy of the Primary Care Assessment Tool-Child Edition (PCAT-CE) for evaluating the attainment of the key characteristics of primary care services for children and youth.

Design. Community-based telephone survey.

Setting. Specific political subdivision in Washington, DC.

Participants. Four hundred fifty parents/guardians of offspring 18 years of age or less.

Measures. Reliability, validity and principal component analysis of 5 scales representing key aspects of the 4 cardinal domains of primary care included in the PCAT-CE. In addition, 2 subdomains (first contact use and extent of affiliation with a primary care source) were included as indices to describe overall patterns of use and affiliation with the particular source of care.

Results. Most scales had adequate internal consistency, test-retest reliability, and construct validity. The principal components factor analysis yielded 5 separate factors. These corresponded to the subdomains of first contact accessibility; coordination of care; characteristics of the professional-patient relationship over-time; and comprehensiveness (both services available and indicated services received).

Conclusions. Psychometric assessment supported the integrity and general adequacy of the PCAT-CE for assessing the characteristics and quality of primary care for children and youth. Testing of revised versions in a variety of different settings is underway. A major component of this testing is to explore the possibility of reducing the number of items while retaining sufficient detail about each component of primary care to make judgements about people’s experiences with that care.


ABBR EVIATIONS. IOM, Institute of Medicine; SCHIP, State Children’s Health Insurance Program; PCAT-CE, Primary Care Assessment Tool-Child Edition.

Efforts to improve the accountability of health service organizations are likely to increase over time. An accumulating literature documents the importance of a strong primary care infrastructure within a health services system. Health systems that are better organized around a primary care base achieve better health outcomes and greater population satisfaction at lower costs than health systems more focused on specialty services.1 Within the United States alone, areas with more primary care physicians have lower rates of hospitalization for conditions that are sensitive to primary care,2 lower death rates from major causes,3 and better birth outcomes.3 A recent report by the Institute of Medicine (IOM)4 reinforces the importance of strong primary care as a critical feature of the health services system. It also encouraged the development of appropriate tools to assess its attainment. The challenge is further heightened by the development of new programs, such as the State Children’s Health Insurance Program (SCHIP), which has made federal funds available to states to provide health insurance for low-income children.

Characteristics of health services delivery that are unique to primary care are theoretically amenable to measurement.5 As a result, it is now possible to assess health services delivery systems according to the characteristics of their approach to providing primary care. Because primary care services should be person-focused rather than illness- or problem-focused, the assessment of the quality of these services is appropriate to all individuals and populations.

A previously described model of primary care,5 which applies to all ages, was used as the framework to guide the development of a Primary Care Assessment Tool-Child Edition (PCAT-CE). Over the past 30 years, a succession of professional committees and experts have reaffirmed the important characteristics of primary care, although the terms that are used to describe these are not always the same. Taking direction from one of the first official documents on the subject,6 we have designated these characteristics as: first-contact care (facilitating entry into care), longitudinality (ongoing person-focused care), comprehensiveness, and coordination. All have been documented to be important in primary care as a means of reducing the likelihood of unnecessary and potentially dangerous care as well as reducing overall costs of care.5 These 4 domains, in combination, are considered to represent the cardinal features of primary care. Each of these domains encompasses 2 subdomains: the capacity of the provider’s practice (structure) and the activities of the practice that pertain to the characteristics of primary care (process).5

Our approach is intended to describe people’s
actual experiences in interacting with the health system, rather than their perceptions of its adequacy (satisfaction). Because experiences have been shown to relate to satisfaction,7 a focus on the respondent’s experiences should also reflect satisfaction. Thus, the research sought to determine if the PCAT-CE is a potentially useful tool for assessing the adequacy of key characteristics of primary care services for children and youth, and to explore the psychometric properties of the pilot instrument.

PARTICIPANTS AND METHODS

Participants

The city of Washington, DC was chosen as the site for testing the instrument because the health department’s strategic plan included an assessment of primary care. The city was particularly interested in health services for children and youth and requested that the survey be targeted at primary care and be conducted in a political subdivision that represented a cross-section of the city’s population. The selected ward had a population of 72,118, with 12,961, or 18%,<18 years old.8 A national survey company assisted in the development of a sampling frame using random digit dialing and telephone listings. A target sample size of 450 respondents was defined to achieve 90% power to detect differences between relatively deprived and nondeprived population subgroups with an overall expected proportion of 50%, assuming a binomial distribution, a significance level of .05 and a 20% nonresponse rate.9 The survey instrument and procedures were approved by the Johns Hopkins University Office for Research Subjects and the District of Columbia Institutional Review Board for Public Health.

Survey Procedures

Cost considerations dictated administration by telephone rather than through in-person interviews. Interviewers (nursing students, graduate students in public health and community members who had been active in community promotion projects) were trained in interactive sessions using a specially designed training manual.

The focus of the interview was selected in each household by asking for the name of the child with the next birthday. The parent/guardian respondent was selected by asking for the person who could best answer questions about getting health care for that particular child. In addition to assessing the characteristics of primary care, the questionnaire ascertained sociodemographic and insurance data as well as information on overall health status. Respondents were not compensated for their participation.

Response rates were based on the number of completed surveys divided by the number of telephone numbers with potential respondents, excluding numbers that corresponded to businesses, fax machines, homes without children <18 years old, disconnected phones, answering machines, busy signals, and no answers. Four phone contacts were attempted for each phone number, using different hours (morning, afternoon, and evenings on weekdays). One third (33.2%) of all contacts were successful and resulted in 450 completed surveys. One fourth (24%) of all phone calls were intercepted by an answering machine. The 64% nonresponse rate is attributed to the following: a) 34% refused to participate in the interview, b) 15% of the interviews were terminated before completion, and c) 15% unspecified reasons. A subsample of 126 respondents selected using random numbers was asked to complete a second, shorter survey within 2 weeks; 35 completed surveys provided data for reliability of responses.

On average, the survey required 25 minutes to administer by telephone. Each questionnaire was independently coded by 2 people (with adjudication by a research team member if necessary) and responses were entered into a computer database by a third person. Data entry accuracy was checked by randomly selecting questionnaires.

MEASURES

Identification of Primary Care Source (Longitudinality: Extent of Affiliation)

Three questions were developed to identify the child’s most likely or usual source of primary care, rather than the “regular” source of care, which might be a specialist consulted frequently for care of some particular problem. The respondent was asked: 1) if there was a particular place or doctor that the child usually goes to if sick or when advice about health is needed, and its identity (“usual source”); 2) the identity of the place/doctor that knows the child best as a person (“knows best”); and 3) the identity of the place/doctor where the child would be taken in the event of a completely new health problem, aside from an emergency (“new problem”). An algorithm based on responses to the 3 questions identified the most likely source of services: care: 1) If all 3 places/doctors were the same, that site was used as the focus of the interview questions; 2) if the response to “usual source” was the same as for either of the other 2 questions, then the “usual source” site was used; 3) if the response for “usual source” was different, but the responses to the 2 other questions were the same, then the place/doctor where both are the same was used; and 4) if all 3 responses were different, then the site identified for “usual source” was used. All subsequent questions in the interview asked about this specific place or person.

Cardinal Domains of Primary Care

As noted in the preceding paragraph, the structural characteristic of actually having a primary care source (Longitudinality: Extent of Affiliation) was assessed by means of 3 questions. Interorough questions concerned the characteristics of the relationship with this source (Longitudinality: Relationship). To assess the structural characteristics that facilitate First-Contact: Accessibility of that regular source of care, 11 questions common in previous studies of accessibility to health care were used. The subdomain of First-Contact: Use concerned the consistency of use of that source when care was last needed for each of 3 different types of need. The Comprehensive-Coordination: Services Provided subdomain included 5 items that asked about the actual receipt of indicated age-specific services. The domain of Coordination addressed only the actual integration of services between a primary care provider and specialty care. The “Don’t Know” option was included for respondents who did not know the characteristics of the practice (structure) that facilitate coordination of care. These items were asked only if the referent child ever had a visit to a specialist or a special service.

SCORING

Scales that were tested by standard psychometric techniques represented all but 2 subdomains for the 4 cardinal characteristics of primary care. The sum score for each scale was derived by scoring each response category to indicate how certain the respondent was about the specific service provided by the primary care provider.

Cerainty as to whether a service was available or provided was scored as a Likert scale with scores of “1” for “Definitely Not”, “2” for “Probably Not”, “3” for “Probably”, and “4” for a “Definitely” response. An additional “Don’t Know” option was included for respondents who answered in this manner. Because a “Don’t Know” response represents a lack of knowledge about a characteristic that should be known to people using a specific primary care site, it was scored as 0. The responses to questions in each subdomain or domain were averaged to obtain the subdomain or domain score.

It was considered to be indices rather than scales because the individual’s experiences for these aspects of care are best represented by a sum of item responses rather than an average of the responses for the different items. Thus, for these 2 subdomains, standard psychometric testing was not used. The index for the subdomain of First-Contact Care: Use was derived from responses to the 3 questions concerning the place/doctor where care was sought the last time it was needed for each of 3 different types of needs (regular check-up, immunization, and illness). A score of 4 was assigned if the same place/doctor was used for all 3 types of needs; a score of 3 was
assigned if 2 of the 3 places/doctors were the same; and if 3 different places/doctors were identified for each need, a score of 0 was assigned. The index for the subdomain of Longitudinality: Extent of Affiliation was assigned according to the responses on the 3 questions noted above regarding the likely source of primary care. A score of 4 was assigned if the child’s regular place (or doctor) of care (“usual place”), place knowing the child best (“knows best”), and place that care would be sought for new problem were the same (“new problem”). A score of 3 was assigned if 2 of the 3 places were the same; and if all 3 were different, a score of 0 was assigned.

ANALYSIS

Responses to the survey questions were used to test the reliability and validity of the instrument. A large number of items (>100) were originally included in the test instrument. The final subdomain items were selected based on various aspects of validity and reliability. In addition, a high percentage of “Don’t Know” responses were used to eliminate items from the final scale.

Validity Assessment
Validity refers to the extent to which an inference is appropriate or meaningful.10 Two aspects of validity were assessed: content validity and construct validity. Content validity is the extent to which the scale is appropriate to its intended purpose and whether it adequately reflects the intended purpose.10 Construct validity is achieved when evidence supports a proposed interpretation based on theoretical implications; this evidence consisted of findings on confirmatory factor analysis and item-total correlations.

Content Validity10

Nine experts were asked to rate the appropriateness and representativeness of the primary care domain items. These experts consisted of 3 policymakers in federal agencies; 2 directors of community pediatrics at major medical centers; 1 health research director at a major health maintenance organization; 2 family medicine professors; and 1 general internal medicine physician with expertise in primary care.

A “card sorting technique” was used to determine the degree of congruence between each item and the domain it was designed to measure. Each survey question with its response categories and descriptions of each of the primary care domains were printed on separate index cards and mailed to the experts who assigned each question to one of the defined domains, and suggested revisions and/or addition of other items. The percent agreement among the experts was used to determine the degree of congruence on the placement of each item in a particular domain. In addition, students in a graduate course on primary care independently assigned each item to a domain as well as to its appropriate subdomain.

Construct Validity10

Partial evidence concerning construct validity was obtained by means of a principal components analysis to examine the relationship among the survey items and to see if the pattern of results could be explained by a smaller number of underlying constructs regarding primary care. The analysis included the 33 initial PCAT-CE survey items tested for inclusion in the primary care scales. Responses to these items were used to examine if they confirmed the hypothesized structure with respect to the 5 scales and specific items contained therein. Five components were extracted corresponding to the hypothesized number of scales. Factor loadings were considered significant if they were .40 or higher.13 The components were tested for correlation and in its absence an orthogonal or uncorrelated method of rotation was used.

Reliability Assessment

Reliability is the extent to which scores are consistent, dependable, or repeatable.10 Two estimation procedures were used to determine reliability: test-retest and internal consistency.10 To reduce respondent burden, test-retest only addressed 4 of the 5 primary care subdomains. Longitudinality: Extent of Affiliation, which elicited the identity of the primary care source was not repeated because we wanted to assess reliability of experiences with the same place or doctor.

Given the categorical structure of many of the items and the fact that repeated measures were necessary, 3 indices of item stability were computed: 1) percentage of absolute agreement (A), 2) the kappa (κ) coefficient, and (3) pi (P). These 3 indices were used because they assess different characteristics of stability: an absolute measure (A) and a measure reflecting the ratio of subject to total variation (K). Because the value of K alone is highly related to prevalence and may be low even if the value of A is high, Chinn and Burney12 recommend the use of pi, which assumes zero true prevalence. Ideally, both kappa and pi should be near 1, but items with a low K value should not be rejected unless the value of pi is also low. The criterion of 50% or above for pi was used to inform item selection criteria for the final scale composition.

To determine internal consistency reliability, coefficient α13 and item-total correlations for items in each domain and subdomain scale and inter-item correlations were computed.

RESULTS

The characteristics of the respondent sample were similar to those of the target population as reflected in census data. Race and household income for the survey sample and 1990 census data were as follows, respectively: for race, black (76% vs 65%); white (15% vs 33%); Hispanic (1% vs 2%), and other races (9% vs 3%). Household income for the survey sample and 1990 census data, respectively were $0 to $9999 (23% vs 18%); $10 000 to $14 999 (6% vs 6%); $15 000 to $24 999 (9% vs 16%); $25 000 to $34 999 (10% vs 14%); and over $35 000 (33% vs 46%).8

Although the response rate was lower than that usually desired in studies seeking to generalize from samples to populations, it is less important in studies that seek to test psychometric properties of instruments. The ultimate test of the psychometric soundness of an instrument is replicability in different population groups, not high response rates. In contrast, high response rates are important when the aim is to accurately describe a population’s experiences with primary care of different types.

PSYCHOMETRIC CHARACTERISTICS

Table 1 presents the psychometric characteristics of the 5 primary care scales. They include: 1) percent agreement regarding placement in the a priori hypothesized domain among the panel of experts (content validity); 2) item-total correlation (r) (construct validity); 3) percent “Don’t Know” responses; 4) test-retest reliability (percent absolute agreement; kappa; pi); 5) and internal consistency reliability (α coefficients) of the final set of items selected for each of the primary care subdomains measured. The following summarizes the properties of each of the subdomains in the order of: content validity, construct validity, test-retest, and internal consistency reliability.

LONGITUDINALITY: EXTENT OF AFFILIATION INDEX

Two-thirds (n = 302) of the respondents identified the same provider or place for the 3 questions regarding “regular” source, the source knowing child best, and the likely source of care for a new problem. One fourth (n = 112) identified the same
place or doctor for 2 of the 3 questions; and 8% (n = 36) identified 3 different places/doctors. “Don’t Know” and missing responses accounted for 3% of the responses for name of regular place; 6% for both place/doctor knowing person best, and 6% for place/doctor where care would be sought for a new problem. Because these questions resemble but expand on questions used in surveys that elicit people’s regular source of care, the expert panel was not asked to assign the questions to a domain.

LONGITUDINALITY: RELATIONSHIP SCALE

Table 1 displays the 4 items selected to measure the subdomain of Longitudinality: Relationship. The level of agreement among the expert panel on the items ranged from 11% to 100%. Test-retest results ranged from 54% to 89% for absolute agreement. Internal consistency for the 4 items selected to represent the subdomain of Longitudinality: Relationship was an α coefficient of .40. Because these results were considered inadequate, new items have been added to a revised instrument.

FIRST-CONTACT: ACCESSIBILITY SCALE

Table 1 displays the 4 items selected to represent the subdomain of First-Contact: Accessibility. Expert panel judgments ranged from 78% to 100% agreement on their relevance to the concept of First-Contact. Test-retest results ranged from 37% to 69% for absolute agreement. Internal consistency for the 4 items selected to measure First-Contact: Accessibility was an α coefficient of .68.

<table>
<thead>
<tr>
<th>TABLE 1. Psychometric Properties of Primary Care Scales</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain/Item</td>
</tr>
<tr>
<td>--------------</td>
</tr>
<tr>
<td>--------------</td>
</tr>
<tr>
<td>Longitudinality-relationship (alpha = .40)</td>
</tr>
<tr>
<td>Seen by same provider</td>
</tr>
<tr>
<td>Provider understands patient</td>
</tr>
<tr>
<td>Provider answers questions the way patient understands</td>
</tr>
<tr>
<td>Have regular phone hours</td>
</tr>
<tr>
<td>First contact accessibility (alpha = .68)</td>
</tr>
<tr>
<td>Open weekends</td>
</tr>
<tr>
<td>Open evenings</td>
</tr>
<tr>
<td>Seen weekends, if closed</td>
</tr>
<tr>
<td>Seen same night, if closed</td>
</tr>
<tr>
<td>Comprehensiveness-services available (alpha = .79)</td>
</tr>
<tr>
<td>Social services</td>
</tr>
<tr>
<td>WIC services</td>
</tr>
<tr>
<td>Dental check-ups</td>
</tr>
<tr>
<td>Dental treatment</td>
</tr>
<tr>
<td>Family planning</td>
</tr>
<tr>
<td>Drug/alcohol counseling</td>
</tr>
<tr>
<td>Mental health counseling</td>
</tr>
<tr>
<td>Comprehensiveness-services provided (alpha = .81)</td>
</tr>
<tr>
<td>Discussion of Healthy behaviors</td>
</tr>
<tr>
<td>Personal safety</td>
</tr>
<tr>
<td>Home safety</td>
</tr>
<tr>
<td>Behavior problems</td>
</tr>
<tr>
<td>Growth changes</td>
</tr>
<tr>
<td>Coordination (alpha = .86)</td>
</tr>
<tr>
<td>Suggested taking child to specialist</td>
</tr>
<tr>
<td>How did you find specialist</td>
</tr>
<tr>
<td>Knows patient going to specialist</td>
</tr>
<tr>
<td>Who made appointment for specialist</td>
</tr>
<tr>
<td>Composite score; Information written re: visit</td>
</tr>
<tr>
<td>Talked with you after visit to specialist</td>
</tr>
</tbody>
</table>

* The value of kappa and pi cannot be calculated when the absolute agreement data forms a rectangular matrix.
COMPREHENSIVENESS: SERVICES AVAILABLE SCALE

All 12 items originally designed to measure comprehensiveness of services available were placed in this domain by 100% of the expert panel. Test-retest results ranged from .46% to 71% for absolute agreement. The 7 items retained for the subdomain of Comprehensiveness: Services Available had an α coefficient of .79.

COMPREHENSIVENESS: SERVICES PROVIDED

All 5 of the items were considered to represent comprehensiveness by 8 of the 9 experts. Test-retest findings were moderate (.43–.71) for the items, with the lowest absolute agreements being for discussion of personal safety (.43) and home safety (.51). The subdomain scale for Comprehensiveness: Services Provided had an α coefficient of .81.

COORDINATION: INTEGRATION SCALE

Four of the 6 items were considered to represent coordination by 89% of the experts. Test-retest results were moderate for most items, ranging from 56% to 83%. The 6-item Coordination: Integration scale had an α coefficient of .86.

PRINCIPAL COMPONENTS ANALYSIS

The 5 rotated components extracted explained 48% of the common variance in the responses to the original 33 primary care scale items in the PCAT-CE (rotated components matrix available from the authors). The rotated factors displayed little or no correlation so an orthogonal rotation method was selected.

Taking into consideration only significant component loadings (> .40), the resulting rotated component matrix reflected a simple structure. That is, most items loaded significantly on only 1 of the 5 components representing the 5 hypothesized primary care scales. The correspondence between the hypothesized survey items and the ones that loaded on each component was generally good but varied according to the primary care dimension. The items that loaded most highly on Component 1 were the same 6 items that measured the Coordination: Integration subdomain. Component 2 includes all 5 items in the Comprehensiveness: Services Received subdomain as well as an additional item regarding lead poisoning tests from the other comprehensiveness subdomain. (This item was eliminated from both final scales.) Component 3 included 7 of the 11 original items in the Comprehensiveness: Services Available subdomain. Component 4 includes 4 of the 15 original items hypothesized to measure First Contact: Accessibility. Finally, Component 5 corresponded to all 4 items in the Longitudinality: Relationship subdomain.

COMMENT

The potential benefits of a reliable and valid measure of the achievement of primary care are many. The IOM’s report on primary care4 recognized the importance of studying the potential impact of changes in the organization and financing of health care on primary care within the context of cost containment by means of periodic surveys of individuals of all ages. Because the IOM report did not suggest ways to measure the concepts included in its definition of primary care, they have been interpreted differently by researchers who have developed survey instruments.14–16 Evidence from the principal components analysis reported above suggests that the scales of PCAT-CE measure 5 distinct aspects of primary care; these, with the 2 indices, represent the seven of the 8 subdivisions of the 4 primary care domains that could be judged by consumers (First-Contact, Longitudinality, Comprehensiveness and Coordination). (The Eighth Coordination: Medical Record adequacy is best obtained by assessing the practice itself).17

Because some additional domains (Family-Centeredness, Community-Oriented, and Cultural Competence) are sometimes considered critical to primary care,4,18,19 we also tested the performance of relevant items. Analysis (not reported herein) indicated that these domains were related to the cardinal domains, thus supporting the previous conceptualization9 of these characteristics as subsidiary or “derivative.” Because of this and the borderline adequacy of their psychometric properties, they are considered supplementary, rather than integral, to primary care assessment.

The number of questions concerning the characteristics of primary care has been reduced by more than half. Because the questionnaire also contained questions on health insurance, sociodemographic characteristics, the type of place where care is sought and its characteristics, none of which are needed to characterize primary care itself, the total reduction in questionnaire length is about 25%. Reduction in the length of the instrument should facilitate administration and reduce respondent burden.

The applicability of the Primary Care Assessment Tool to assess the adequacy of primary care experiences is promising. Additional analyses of these data as well as data from a comparison provider survey demonstrated differences in the adequacy of primary care between managed care and more conventional forms of practice.17 The survey in its original form has been used to compare 4 different types of managed care plans in Florida’s Healthy Kids program,20 and a version modified for adults is now available. Testing of both adult and child versions is underway in several clinical sites, including community health center and health maintenance organization settings. A self-administered version (as for administration by mail) has also been prepared. Repeated analyses in these different settings and under different conditions should help to determine the extent to which the instrument’s psychometric properties are consistent across different population subgroups.

The results of this study as well as others currently underway suggest that the instrument, particularly those parts addressing the cardinal do-
ACKNOWLEDGMENTS

The Consumer Primary Care Assessment Tool, was produced by the Child and Adolescent Health Policy Center, Department of Maternal and Child Health, Johns Hopkins University School of Hygiene and Public Health. This work was supported by a Federal Set-Aside Grant No. MCJ243A19 from the Maternal and Child Health Bureau (Title V, Social Security Act), Health Resources and Services Administration, Department of Health and Human Services, and a grant from the Henry J. Kaiser Family Foundation.

Field tests of the Consumer Primary Care Assessment Tool, were supported by the District of Columbia’s Office of Maternal and Child Health through a Federal Set-Aside Grant No. MCJ42B032, “Development of a Comprehensive Primary Care System of Services for All Children,” from the Maternal and Child Health Bureau (Title V, Social Security Act), Health Resources and Services Administration, Department of Health and Human Services.

Special appreciation is extended to the following people whose time, expertise, and commitment was invaluable in this project: Pat Tompkins, MPH, Chief (retired), Office of Maternal and Child Health, DC Commission of Public Health; Barbara Hatcher, PhD, Chief, Office of Maternal and Child Health, DC Commission of Public Health; and Joni Eisenberg, Project Coordinator, Office of Maternal and Child Health, DC Commission of Public Health.

REFERENCES