Quality and Satisfaction

Women's Satisfaction with Their On-Going Primary Health Care Services: A Consideration of Visit-Specific and Period Assessments

Roger T. Anderson, Carol S. Weisman, Fabian Camacho, Sarah Hudson Scholle, Jillian T. Henderson, and Deborah F. Farmer

Objective. To compare and contrast patient ratings of satisfaction with primary care on the day of visit versus over the last 12 months.

Data Sources/Study Setting. Survey data were collected from female participants at primary care centers affiliated with the University of Michigan, University of Pittsburgh, and Wake Forest University.

Study Design. One thousand and twenty-one patients attending a primary care visit with at least one prior visit to the study site were consented on site, enrolled in the study, and surveyed at two time points: pre- and immediately postvisit.

Data Collection. The previsit survey included demographics, self-rated health, visit history (site continuity), and expectations for health care; the postvisit survey focused on patient experiences during the visit, assessment of health care quality using the Primary Care Satisfaction Survey for Women instrument, and global satisfaction with visit and health care over the past 12 months. Expectation discrepancy scores were constructed from the linked expectation–experience ratings. Path analysis and indices of model fit were used to investigate the strength of theoretical links among the variables in an analytic model considering both day-of-visit and past-year ratings with global measures of patient satisfaction as the dependent variables.

Principal Findings. General health, site continuity and fulfillment of patient expectations for care were linked to global ratings of satisfaction through effects on communication, care coordination, and office staff and administration. Importantly, past-year ratings were mediated largely by care coordination and continuity; day-of-visit ratings were mediated by communication.

Conclusion. Ratings of health care quality for a specific visit appear to be conceptually distinct from ratings of care over the past 12 months, and thus are not interchangeable.

Key Words. Satisfaction, primary care, women's health, path analyses

Patient satisfaction ratings are routinely collected to assess the quality of primary care and to assist consumers in selecting a primary care provider. Many different tools for assessing patient satisfaction with primary care are available, and much previous research has addressed the aspects of patient experiences that provide the basis for ratings of satisfaction (Linder-Pelz 1982; Brody et al. 1989; Strasser et al. 1993; Peck et al. 2001; Ford, Schofield, and Hope 2003; Scholle et al. 2004). However, there is less clarity in the literature about the role of the time frame for which the rating is made. Current patient satisfaction rating tools ask respondents to rate their health care experiences either for a specific health care visit or for care received over a period of time (e.g., the past 12 months). For purposes of assessing quality of care within a health plan or practice, these different time frames are not likely to be interchangeable. Some aspects of care, such as the quality of communication between patient and provider or the thoroughness of an examination, may be readily assessed for a specific patient visit; other aspects of care, such as trust in the provider or the quality of decision making, are based on cumulative experiences (Peck et al. 2001) and are likely to be evaluated across several encounters or within a treatment period or "episode" of care (Hornbrook, Hurtado, and Johnson 1985). Visit-specific measures are likely to miss issues related to continuity or coordination of care over time, whereas period ratings may miss issues related to interactions with a specific provider or service. Depending on the purpose for which patient satisfaction ratings are obtained, the time frame for the ratings could be important.

The distinction between the contexts of a specific visit and on-going care may be particularly important for women, who make more primary care visits

Address correspondence to Roger T. Anderson, Ph.D., Professor, Department of Social Sciences and Health Policy, Division of Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, NC 27157. Carol S. Weisman, Ph.D., Professor, is with the Health Evaluation Sciences and Obstetrics and Gynecology, Chief, Division of Health Services Research, Director, Central PA Center of Excellence for Research on Pregnancy Outcomes, Pennsylvania State University, College of Medicine, Hershey, PA. Fabian Camacho, M.S., Research Associate, is with the Department of Social Sciences and Health Policy, Division of Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, NC. Sarah Hudson Scholle, Ph.D., Assistant Vice President, is with the Research and Analysis, National Committee for Quality Assurance (NCQA), Washington, DC. Jillian T. Henderson, Ph.D., is with the Center for Reproductive Health Research and Policy, University of California, San Francisco, CA and also with the Health Research and Policy, University of California, San Francisco, San Francisco, CA. Deborah Farmer, Ph.D., Research Associate, is with the Department of Social Sciences and Health Policy, Division of Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, NC.

than men (Brett and Burt 2001) and who often may see more than one provider (e.g., a generalist and reproductive health care provider) over multiple visits (Weisman and Henderson 2001) to receive comprehensive, age-appropriate clinical preventive services (Weisman 1996; Gallagher, Geling, and Comite 2001; Weisman and Henderson 2001; Henderson, Weisman, and Grason 2002). However, this pattern of care could also imply more costs to the patient, redundancies in services, increased false positives associated with screening, or contradictory advice across visits. Thus, the impact of on going patterns of care on ratings of healthy care quality is largely unknown.

We studied the uniqueness of patient appraisals of satisfaction with on-going and visit-specific care in a multicenter sample of 1,202 women making routine primary care visits. We hypothesized that aspects of primary care conceptualized as visit focused, such as doctor's communication and demeanor and performance of the office staff, would correlate most highly with day-of-visit ratings; whereas ongoing care processes (e.g., follow-up on test results and care coordination) would correlate most highly with ratings of care over the past year.

METHODS

Subjects

Data are from a multisite study of women's satisfaction with primary care, including 1,202 women aged 18 years and older, English-speaking, without apparent cognitive impairment, able to complete the questionnaire without assistance, and attending a primary care visit at the time of the survey. A "primary care visit" included a routine checkup, gynecological exam, prenatal care, acute care, or routine follow-up care with a doctor or other independent practitioner (advanced practice nurse or physician assistant). Excluded were emergency visits and visits to drop off lab specimens or visits for a single procedure such as a flu shot, allergy shot, or contraception injection. The study sites were affiliated with three primary care centers: the University of Michigan, the University of Pittsburgh, and Wake Forest University. Located in different geographic regions, these centers provide access to diverse primary care settings, patient populations and provider types (i.e., residents, primary care and specialist physicians, and advanced practice nurses or physician assistants). At the University of Michigan, subjects were recruited from three clinical areas (internal medicine, family practice, and obstetrics and gynecology)

in one ambulatory care facility. The University of Pittsburgh sites included a major primary care clinic staffed by faculty and resident physicians, and a large obstetrics and gynecology clinic at Magee-Womens Hospital, serving a population diverse in race/ethnicity and income. At Wake Forest University Baptist Medical Center, Winston-Salem, NC, subjects were recruited from an academic multispecialty practice for faculty and resident physicians and from two freestanding family practice satellite clinics, one of which serves a mostly African American residential area of the city. Survey participation was 69 percent of eligible subjects (1,202/1,742). Reasons for declining participation were not having time to remain after the visit (55 percent of nonrespondents), lack of interest (24 percent), too ill to participate (14 percent), and no reason given (7 percent). Included in the present study analysis were 1,021 patients reporting one or more previous primary health care visits in the past 12 months. Institutional Review Board approval of the study protocol was obtained at each study site.

Study Design

Because this study was focused on comparing ratings of health care quality for a specific visit and those reflected over a 12-month period of time, the analytic sample was limited to the 1,021 participants of the original sample of 1,202 subjects who reported at least one visit to the site prior to the index visit. This was done to ensure that ratings of care over the last 12 months reflected health care visits beyond the day of survey. Data from all three primary care centers were pooled for analysis.

All survey data for this study were collected on site during the patients' office visit. Upon check-in, patients were invited to participate in the study on a "next available" basis, before seeing their provider. Participants completed a consent form and previsit survey on demographics, visit history, reasons for present visit, and expectations for this visit. Participants also completed a postvisit survey on services received during the visit, satisfaction with care and visit, whether care met expectations, and additional demographic items. Upon completion of both questionnaires, subjects were compensated in cash (\$20) or equivalent-value coupons from local retail vendors.

Measures

Global Satisfaction. A summary rating of patient satisfaction with care "over the past year" was obtained using the CAHPS^{®1} global item (National Committee for Quality Assurance 1998) asking respondents to rate "all of

your health care in the last 12 months from all doctors and other health professionals at this office or clinic” on a scale from 0 (worst health care possible) to 10 (best health care possible). Rating day-of-visit care (Overall Visit Rating) was obtained using a single-item from the MOS Visit Satisfaction Scale (Davies and Ware 1991). Respondents rated their satisfaction with the “overall quality of care at this visit,” using a five-point response set (1 = “not at all satisfied” to 5 = “extremely satisfied”).

Exogenous Variables. Because previous research has shown that ratings of health care quality can be influenced by various patient characteristics as a predisposition or need for a level of treatment, care coordination across visits, and care expectations or preferences (Hall and Dornan 1990; Marshall, Hays, and Rand 1996; O'Malley, Forrest, and O'Malley 2000; Rao et al. 2004), included in the initial model were: age group (<45, 45–54, 55–64, 65+ years), race/ethnicity (white, black, and other), educational attainment (high school or less, some college, or college graduate), and perceived general health assessed with a single overall health rating (“excellent” to “poor”) (Ware and Hays 1998). Patient history was modeled as an indicator of potential need for treatment from the number of self-reported diagnoses (yes/no) in the last 5 years involving: hypertension, high cholesterol, heart disease, cancer, diabetes, asthma, depression, migraine headaches, arthritis, osteoporosis, obesity, incontinence or leakage, eating disorders, thyroid problems. To assess whether the relationship of timeframe for the patient rating of health care experience was influenced by indices of care continuity, two measures of continuity were tested: site continuity was included in the initial model as the proportion of all health care visits in the past 12 months made to the clinic (Magill and Senf 1987; Gill and Mainous 1998). Provider continuity was included as the number of regular providers seen for care (labeled: “one provider”) as a two-category variable indicating whether the participant had: one regular primary care provider or not (e.g., either no regular provider or two regular providers). Together, this set of exogenous variables allowed us to explore whether ratings of health care quality for each timeframe were influenced by levels of patient factors such as demographic status, general health, and patterns of health care use. Exogenous variables that do not display statistically significant correlations ($\alpha = 0.05$) with other variables in the initial model were dropped in subsequent models.

To investigate the extent that our patient ratings of health care quality for a specific visit and for the last 12 months were theoretically similar

constructs, representing different aspects of health care, we assessed the discrepancy between expected care and delivered care, labeled here as "expectation discrepancy." Specifically, we expected to find that unfulfilled expectations (negative expectation discrepancy) (Michalos 1985) would lead to low ratings of health care quality, and met or exceeded expectations (zero or positive expectation discrepancy) would lead to high ratings of health care quality. Patient's general expectations for health care were assessed with an independent set of five items using a three-point response scale of importance ("not important," "somewhat important," "very important") before the office visit. The items were developed to parallel the content of the validated Primary Care Satisfaction Survey for Women (PCSSW; Scholle et al. 2004): "talking to the health professional with my clothes on" (privacy); "getting everything I need at one visit" (comprehensiveness of care); "getting help scheduling my next appointment" (coordination of care); "having a health professional who includes me in decisions about my care" (decision making); and "having a health professional who coordinates all the health care I receive" (coordination/continuity). After the visit, patients' rated the extent to which each of these experiences occurred independently of the PCSSW, using a three-point response scale ("alot," "some," "not at all") to assess the degree that each expected experience occurred. An expectations discrepancy score was calculated for each item, ranging from -2 to $+2$, where 0 indicates an exact match of patient visit experiences with stated expectations, $a-2$ indicates large negative dissonance (two response category units below patient expectations), and $a+2$ indicates large positive dissonance (two response category units above expectations). An overall discrepancy score was taken as the mean of the item scores.

Endogenous Variables. Hypothesized mediators of global ratings of satisfaction with primary care are subscales of the PCSSW assessing Communication (eight items, range: 8–40, $\alpha = 0.96$), Office staff and Administrative Procedures (six items, range: 6–30, $\alpha = 0.88$), and Care Coordination and Comprehensiveness (10 items, range: 10–50, $\alpha = 0.85$) with higher scores indicating greater satisfaction (Anderson et al. 2001; Scholle et al. 2004). Items include a focus on both day of visit and past care over a 12-month period such that ratings of communication and office staff are obtained for the present visit, and care coordination is rated based on perceived quality for all visits (including present) over the past year. Each PCSSW item is rated on a five-point scale: 1 = "not at all satisfied," 2 = "somewhat satisfied," 3 = "satisfied," 4 = "very satisfied," and 5 = "extremely satisfied" point scale.

Analytic Model

Path analysis, a form of Structural Equation Modeling (SEM), was used to compare the strengths of hypothesized direct and indirect relationships (correlations) among a set of variables in cross-sectional data (Wright 1934; Davis 1985; Asher 1988). Investigated were the causal links between exogenous variables (including: patient age, race/ethnicity, provider and site continuity, expectations, and general health), endogenous variables modeled as mediators (PCSSW scales), and endogenous variables as the terminus in the model (global satisfaction ratings with present visit and care over the past 12 months).

Model fit was tested by comparing how well the estimated correlation matrix under the model approximates the observed correlation matrix. Diagrams illustrate causal connections between the variables as straight unidirectional arrows, co-variation between variables as double-headed arrows, and errors as latent random variables. Each of the connections is associated with a regression weight, and a variance-covariance matrix is derived to test for model fit against the sample variance-covariance matrix (Pedhazur 1982; Davis 1985; Asher 1988; Hatcher 1988) using a maximum likelihood method. Since many of the variables may not be normally distributed, bootstrapping methods were used to obtain robust standard error estimates and confidence intervals. All analyses were performed on the variance-covariance matrix. The structural equation modeling program AMOS was used to fit the model and estimate model parameters (Arbuckle and Wothke 1999). Approximately 5 percent ($N = 52$) of the study sample had incomplete survey data, with a maximum frequency of four missing item responses. In order to retain these observations in the analyses, missing data were imputed by using a Markov chain Monte Carlo method (Schafer 1997) and using all numerical predictors available.

Since the paths among the three tiers of variables (i.e., exogenous, endogenous, and global satisfaction) described above were not completely specified a priori, some exploratory modifications were performed in order to develop plausible models with the best trade-offs in parsimony, substantive interpretation, and goodness of fit (Joreskog and Sorbom 1984). Indices of goodness of fit to evaluate the models were: the chi-square statistic as a test of the null hypothesis that the sample covariance matrix stems from the model (expected to be $p > .05$ due to large sample size); and the comparative fit index (CFI) which assesses the congruence between model and data (Hu and Bentler 1995). The CFI ranges from 0 to 1, where 0 represents the fit in the null model in which all variables are modeled as uncorrelated, and 1 represents the fit of

the saturated model in which enough parameters exist to replicate the sample covariance matrix without error. According to Bentler and Bonett (1980) models with a CFI < 0.9 can be substantially improved. Thus, a value of 0.9 or greater is commonly regarded as indicating excellent goodness of fit.

Because exploratory work was performed in order to find the best fitting models, it was necessary to use indices and methods that adjusted the goodness of fit for model complexity. These included the Akaike information criterion (AIC) (Akaike 1987), the expected cross-validation index (ECVI), the minimum discrepancy score divided by the number of degrees of freedom (Cmin/df), and the root mean square error of approximation (RMSEA). A ratio of 2–3 for the minimum discrepancy score per degrees of freedom indicates an acceptable fit according to Carmines and McIver (1981), and a value of < 0.05 indicates a close fit of the model, with 0.08 or less being reasonable for the RMSEA (Cudeck and Browne 1983). For the AIC and ECVI, goodness of fit is compared against the saturated model. A boot-strapping procedure as described in Linhart and Zucchini (1986) and Arbuckle and Wothke (1999) also was performed to compare the fit between models.

Model stability was checked across clinic site, age categories, and race/ethnic categories with multiple-group analysis (Lee and Tsui 1982; Arbuckle and Wothke 1999) using a likelihood ratio test of constrained and unconstrained models to judge whether the constrained model does not significantly change the goodness of fit. Nonsignificant results imply nonheterogeneity in the causal weights across different subgroups. Weights deemed to be heterogeneous are allowed to vary. Models resulting from lesser constraints are compared again until fit is judged not to significantly deteriorate from the unconstrained model.

RESULTS

Table 1 presents the characteristics of the study sample ($N = 1,021$). Approximately 27–30 percent of respondents were in each age group of: < 30 , 30–44, or 45–64 years, and 14 percent were 65 years and older (mean age of 43 years). More than two-thirds (67.7 percent) of were white (non-Hispanic), 22.5 percent were black (non-Hispanic), and 9.8 percent of other race/ethnicity. Education and household income varied widely; 31 percent of women had a high school education or less, 31 percent reported some college, and 39 percent reported having a college degree. Most had an average of 5.4 visits/year (standard deviation = 5.6), with the focus of current visit for a routine exam

Table 1: Descriptive Characteristics of the Study Sample ($n = 1,021$)

<i>Characteristic</i>	<i>N</i>	<i>%</i>
Age (years)		
<30	284	27.93
30–<45	307	30.19
45–<65	282	27.73
65+	144	14.16
Race/ethnicity		
White, non-Hispanic	691	67.68
Black, non-Hispanic	230	22.53
Other/multiethnic*	100	9.80
Education		
High school or less	307	30.52
Some college	308	30.62
College graduate+	391	38.87
General health (SF-36) [†]		
Excellent	129	12.66
Very good	359	35.23
Good/fair/poor	531	52.11
Number of health conditions		
0	258	25.27
1	207	20.27
2	154	15.08
3+	402	39.37
Expectations discrepancy score [‡]		
0	91	89.13
1	69	6.76
2	42	4.12
Site continuity [§]	842	82.47
Provider continuity [¶]	400	39.18

*“Other” race/ethnicity includes those with Hispanic ancestry, other racial/ethnic identification, and multiethnic identification.

[†]Single item assessed previsit: “In general, would you say your health is: excellent, very good, good, fair, or poor.”

[‡]Shown as the number of items out of five in which health care experience was rated below expectation.

[§]Patient visited study-site clinic or practice for all her usual health care visits in past 12 months.

[¶]Patient reported seeing one primary care provider for all of her care in past 12 months.

(22.6 percent), prenatal care (15.9 percent), a new health problem (26.3 percent), or follow-up care (34.6 percent). About one-half (52 percent) rated their health as “good,” “fair,” or “poor”; 35 percent as “very good”; and only 13 percent as “excellent.” For care continuity, 82.5 percent reported the study clinical site as their usual source of care; about 39 percent had one regular provider; 61.8 percent had either no or more than one provider. Most

Table 2: Variables Retained in Path Analysis ($n = 1,021$)

<i>Variable</i>	<i>Mean (SD)</i>	<i>[Minimum, Maximum]</i>
Exogenous		
Age (years)	42.76 (16.82)	[18–82]
General health (SF-36)*	2.58 (0.98)	[1—excellent, 5—poor]
Site continuity [†]	0.68 (0.32)	[0—no, 1—yes]
Expectations discrepancy score [‡]	1.05 (0.57)	[–2—unmet, 2—exceeded]
Endogenous		
Communication (PCSSW [§])	33.69 (6.42)	[8 low, 40 high]
Office satisfaction (PCSSW) [¶]	23.02 (4.69)	[6 low, 30 high]
Care coordination (PCSSW)	38.62 (8.22)	[10 low, 50 high]
Outcome		
Overall rating past 12 months (CAHPS)**	8.60 (1.52)	[0—worst, 10—best]
Overall visit rating ^{††}	4.24 (0.85)	[1—poor, 5—excellent]

*Global self-rated health: “excellent” to “poor” (Ware and Hays 1998).

[†]Site continuity is the proportion of all health care visits in the last 12 months to the study site.

[‡]Difference between pre/postvisit expectations for health care qualities.

[§]PCSSW: the Primary Care Satisfaction Survey for Women (Scholle et al. 2004).

[¶]Office Staff and Administrative Procedures Scale.

^{||}Care Coordination and Comprehensiveness Scale.

**CAHPS[®] global item rating of health care quality at clinic in the last 12 months.

^{††}Global rating of health care quality at present visit (Davies and Ware, 1991).

participants had private insurance (62.5 percent), followed by Medicaid (18.6 percent), and Medicare (12.7 percent) (data not shown).

Shown in Table 2 are the variables retained in the path analytic models based on evidence of statistically significant correlation ($p < .05$) with either the endogenous (mediating) variables or global satisfaction outcomes, or both. Excluded on this basis were provider continuity and patient age.

The path analysis model in Figure 1 depicts the exogenous variables using dashed lines and endogenous variables with solid lines. General health and site continuity have small but significant indirect effects on global ratings of care by way of satisfaction with PCSSW scales: Communication and with Care Coordination and Comprehensiveness. Visit discrepancy scores are strongly associated with each PCSSW scale. The PCSSW scale scores are differentially related to the global ratings in the predicted pattern: Care Coordination and Comprehensiveness is most highly correlated with the CAHPS rating of care over the past 12 months, whereas Communication has the highest correlation with the global visit rating. Administration and Office Procedures contribute only to the visit rating. Because indices estimating change in model fit when a new parameter is added suggested the original

Figure 1: Path Analysis Model

Notes: Significant standardized regression weights and correlations are shown. Thick arrows represent magnitudes that are stable across sites.

model with uncorrelated errors was mis-specified (the two global outcomes, ratings for care during the past 12 months and at the visit, may have unknown causes in common) correlated errors were allowed in the final model shown with double-head curves.

Goodness of fit indices suggests that this model explains the sample data adequately. Although, according to the chi-square test, the model is rejected at this sample size ($n = 1,021$), the CFI indice is almost equal to 1 (0.97), indicating excellent relative fit. The RMSEA (0.02) and the Cmin/df (1.2) satisfy the criteria for a close and acceptable fit of the model, respectively. Furthermore, the AIC and the ECVI are close to the values observed in the saturated model.

In Table 3, the strongest standardized regression weights (i.e., above 0.30) are those corresponding to the weights between expectations discrepancy and the Communication scale (0.70), expectations discrepancy and the Care Coordination and Comprehensiveness scale (0.62), expectations discrepancy and the Administration and Office Procedures scale (0.64), Care Coordination and Comprehensiveness and the CAHPS rating (0.54), and Communication and the visit rating (0.46). Weaker, significant weights below 0.30 include those between Care Coordination and Comprehensiveness and the visit rating (0.26), Administration and Office Procedures and the visit rating (0.13), Communication and the CAHPS rating (0.08), site continuity and PCSSW Care Coordination and Comprehensiveness (0.04), and general health and Communication (0.06). The squared multiple correlations indicate that the Communication scale has approximately 49 percent of its variance explained by its predictors, and Care Coordination and Comprehensiveness has 39 percent of its variance explained. The CAHPS rating, Administration and Office Procedures, and the visit rating have 38, 40, and 60 percent of their variances explained, respectively.

No statistically significant differences among regression weights are observed when the model is fitted separately across different race/ethnic subgroups (white, black, and other) or age groups (18–35, 36–54, and 55 years and over). However, some weights had to be allowed to vary freely between the three different sites in order for the overall model not to be rejected by the likelihood ratio test. These weights were those assessing the relationship between expectations discrepancy and Care Coordination and Comprehensiveness (0.54, 0.67, 0.68 for sites 1, 2, and 3, respectively) as well as those between Care Coordination and Comprehensiveness and the CAHPS rating (0.58, 0.41, 0.35, respectively).

DISCUSSION

The results of this study have important practical implications for the measurement of health care quality from the patient point of view. The data strongly suggest that compared with formulating judgments about the specific visit, patients weigh different pieces of information differently when formulating judgments about health care quality over a period of time. Despite the fact that all ratings were collected at the same point in time, when asked to rate care for a specific visit, respondents relied more upon visit-focused aspects of care, such as provider communication, and less on care coordination occurring over past visits at that same location or even by the same provider (the path analytic model presented above was confirmed in the subset of patients

Table 3: Standardized Regression Weights, Correlations, and Squared Multiple Correlations for the Best Path Model

<i>From</i>	<i>To</i>	<i>Standardized Regression Estimate</i>	<i>95%* CI Lower Bound</i>	<i>95% CI Upper Bound</i>	<i>SMC[†]</i>
Expectations discrepancy	Care coordination (PCSSW)	0.621	0.569	0.663	0.387
Site continuity	Care coordination (PCSSW)	0.038	0.001	0.081	
Expectations discrepancy	Office staff/administrative (PCSSW)	0.635	0.594	0.671	0.404
Expectations discrepancy	Communication (PCSSW)	0.698	0.659	0.733	0.490
General health (SF-36)	Communication (PCSSW)	0.055	0.026	0.090	
Care coordination (PCSSW)	Positive rating past 12 months (CAHPS)	0.543	0.462	0.617	0.375
Communication (PCSSW)	Positive rating past 12 months (CAHPS)	0.084	0.011	0.165	
Office staff/administrative (PCSSW)	Overall visit rating	0.127	0.078	0.183	0.602
Communication (PCSSW)	Overall visit rating	0.456	0.376	0.527	
Care coordination (PCSSW)	Overall visit rating	0.261	0.190	0.330	

*Ninety-five percent intervals using the biased corrected percentile bootstrap method (200 subsamples used).

†Square multiple correlations.

who see the same provider for all of their care). Therefore, ratings of overall health care quality for a specific visit are probably not interchangeable or comparable with such ratings that refer to health care quality over a period of time (e.g., the past year). If true, the important implication is that it may not be possible to represent period ratings by averaging over the ratings for specific visits unless the user is interested only in those aspects of quality that are most affected by specific visits. It also means that the quality of health care delivery among practices or organizations should not be compared when the ratings used to describe them vary regarding to the time period specified (e.g., visit specific, versus the past year). As another practical implication, the use of both visit-specific and past-year assessments may provide a more thorough assessment of patient satisfaction than either measure used alone.

Caveats about the breadth of global ratings raised in this study are important for researchers to consider because both time foci are commonplace in generic measures of overall patient satisfaction. One of the authors (D. F.) reviewed all published research reports indexed in Medline under keywords of *patient satisfaction* and *primary care*, published from 1996 through 2002 that included sufficient details on the time reference of the satisfaction appraisal. Of 29 publications identified as meeting our criteria, 10 studies relied solely upon “past-year” ratings, 10 relied solely upon “day-of-visit” ratings, and nine used both past year and day of visit.

The final path analysis model shows that exogenous variables of health status and site continuity exert small but significant effects on rating so of health care quality through communication and care coordination. Persons with lower self-rated health tended to rate provider communication somewhat lower than those with high ratings of health. The latter path is perhaps because poorer health requires more health care interventions and complex treatment decisions and may therefore create greater challenges to clear communication. Site continuity was positively correlated with ratings of care coordination but not with communication or office staff. A finding discrepant with our hypothesized model was that having one versus more than one regular provider was not important predictor of ratings of health care quality made for either timeframe, and was eliminated in the final model as were age group, educational attainment, and race.

Results of this study also supported that the two distinct approaches to global measures of satisfaction appear each to be theoretically founded upon delivery of care that fulfills patient expectations. Patients have expectations or goals for their health care visits and a desire to satisfy those goals is a primary component of quality of care assessment (Bell et al. 2001). Tests of this assumption have yielded mixed results (Cleary and McNeil 1988; Sitzia and

Wood 1997; Peck et al. 2001). We asked patients to assign the importance of various health care expectations prior to the actual visit, and then conducted a postvisit survey to assess expectation fulfillment. The expectations discrepancy scores were strongly associated with ratings of overall satisfaction through each specific PCSSW scale score and indirectly related to overall ratings of care quality for a specific over 12 months. There are no standard methods for assessing patient expectations for health care in the literature, and measurement approaches have varied considerably. The distinction of qualitative and technical aspects of care has been highlighted by Peck et al. (2001) who did not find an association between direct measures of expectations for technical medical services and patient satisfaction in a Veterans Administration primary care clinic, but who hypothesized that patient expectations for nonmedical services would be a more powerful correlate. Our study is consistent with this prediction and with existing literature on the influence of patient-focused expectations for care on satisfaction ratings such as patient participation in decision making (Brody et al. 1989) and perception of provider interpersonal behavior (Froehlich and Welch 1996).

Finally, a limitation of our study is that we studied women patients who had completed at least one additional visit in the past 12 months to the study practice site. This criteria likely selected patients with health conditions that needed periodic management while excluding generally healthy patients with few health concerns. Thus our results cannot be generalized to patients who infrequently accessed the study sites. Another caveat is that we did not collect ratings of communication over the past 12 months and, therefore, could not ascertain the extent to which the latter are important in period ratings of primary care satisfaction. Our study was instead designed to test the validity of the approach taken in the PCSSW of including both care coordination and continuity content and visit-specific information to assess different dimensions of patient care. Asking patients to rate usual levels of communication or performance of office staff and health care providers over several past visits is challenging, as patients may see different providers from visit to visit, and must consider past discussions, possibly over long periods of time, to derive an average rating. For this reason, such ratings are not commonplace in existing measures of patient satisfaction.

The validity of some of the key variables in the model, such as the expectations discrepancy score and the overall visit rating, may be subject to scrutiny. Our measure of expectations discrepancy was constructed for this study without prior validation, to parallel key concepts assessed in validated PCSSW instrument. We were unable to locate similar or suitable measures in

the literature with established validity for this purpose. Not accounting for the actual reliabilities of the variables may affect the estimates in several ways. If the variables have low reliability, the standardized regression weights may be meaningfully attenuated, leading to underestimation of the strength in the relationships (Pedhazur 1982). Attenuated reliabilities on the dependent variables will also result in underestimation of the proportions of variance explained. The magnitudes presented in this analysis are lower bounds of magnitudes that would occur if the variables were to be perfectly reliable. The nonnormality of the variables may also bias the parameter estimates, their standard errors, hypothesis tests, and indices of fit (West, Finch, and Curran 1995). However, the adverse effect of variable skewness on our inferences is minimized because of the large sample size ($n = 1,021$) and our use of bootstrapping methods to calculate robust standard errors. Finally, our study results regarding the PCSSW scales and global ratings are based on cross-sectional data and thus cannot test causality. Instead, we tested our hypothesized model of relationships among variables using path analysis strategies.

CONCLUSION

The results of this study provide evidence that care coordination is more closely related to global ratings of primary care over several visits than are qualities of care experienced at a current visit, where communication is the most salient predictor. This finding suggests that ratings of health care quality for a specific visit are distinct from ratings of care over the past year, and that both types of ratings may be needed to adequately assess the range of core topics relevant to patient satisfaction.

NOTE

1. CAHPS is a registered trademark of the Agency for Healthcare Research and Quality.

REFERENCES

- Akaike, H. 1987. "Factor Analysis and AIC." *Psychometrika* 52: 317–32.
- Anderson, R. T., A. M. Barbara, C. Weisman, S. Hudson Scholle, J. Binko, T. Schneider, K. Freund, and V. Gwinner. 2001. "A Qualitative Analysis of Women's Satisfaction with Primary Care from a Panel of Focus Groups in the National Centers of Excellence in Women's Health." *Journal of Women's Health and Gender-Based Medicine* 10: 637–47.

- Arbuckle, J. L., and W. Wothke. 1999. *Amos 4.0 User's Guide*. Chicago: Smallwaters Corporation.
- Asher, H. B. 1988. *Causal Modeling*. 2d Edition. Beverly Hills, CA: Sage.
- Bell, R. A., R. L. Kravitz, D. Thom, E. Krupat, and R. Azari. 2001. "Unsaid but Not Forgotten: Patients' Unvoiced Desires in Office Visits." *Archives of Internal Medicine* 161: 1977-84.
- Bentler, P. M., and D. G. Bonett. 1980. "Significance Tests and Goodness-of-Fit in the Analysis of Covariance Structures." *Psychological Bulletin* 88: 588-606.
- Brett, K. M., and C. W. Burt. 2001. "Utilization of Ambulatory Medical Care by Women: United States, 1997-98." *Vital and Health Statistics* 13(149):12.
- Brody, D. S., S. M. Miller, C. E. Lerman, D. G. Smith, C. G. Lazaro, and M. J. Blum. 1989. "The Relationship between Patients' Satisfaction with Their Physicians and Perceptions about Interventions They Desired and Received." *Medical Care* 27: 1027-35.
- Carmines, E. G., and J. P. McIver. 1981. "Analyzing Models with Unobserved Variables." In *Social Measurement: Current Issues*, edited by G. W. Bohrnstedt and E. F. Borgatt, p. 80. Beverly Hills, CA: Sage.
- Cleary, P. D., and B. J. McNeil. 1988. "Patient Satisfaction as an Indicator of Quality Care." *Inquiry* 25: 25-36.
- Cudeck, R., and M. W. Browne. 1983. "Cross-Validation of Covariance Structures." *Multivariate Behavioral Research* 18: 147-67.
- Davies, A. R., and J. E. Ware. 1991. *GHAA's Consumer Satisfaction Survey and User's Manual*. Washington, DC: Group Health Association of America.
- Davis, J. A. 1985. "The Logic of Causal Order." In *Quantitative Applications in the Social Sciences Series, No. 55*, pp. 1-72. Thousand Oaks, CA: Sage Publications.
- Ford, S., T. Schofield, and T. Hope. 2003. "What Are the Ingredients for a Successful Evidence-Based Patient Choice Consultation?: A Qualitative Study." *Social Science and Medicine* 56: 589-602.
- Froehlich, G. W., and H. G. Welch. 1996. "Meeting Walk-In Patients' Expectations for Testing." *Journal of General Internal Medicine* 11: 470-4.
- Gallagher, T. C., O. Geling, and F. Comite. 2001. "Use of Multiple Providers for Regular Care and Women's Receipt of Hormone Replacement Therapy Counseling." *Medical Care* 39: 1086-96.
- Gill, J. M., and A. G. Mainous. 1998. "The Role of Provider Continuity in Preventing Hospitalizations." *Archives of Family Medicine* 7: 352-7.
- Hall, J. A., and M. C. Dorman. 1990. "Patient Sociodemographic Characteristics as Predictors of Satisfaction with Medical Care: A Meta-Analysis." *Social Science and Medicine* 30: 811-8.
- Hatcher, L. 1988. *A Step-by-Step Approach to Using the SAS System for Factor Analysis and Structural Equation Modeling*. Cary, NC: SAS Institute Inc.
- Henderson, J. T., C. S. Weisman, and H. Grason. 2002. "Are Two Doctors Better Than One? Women's Physician Use and Appropriate Care." *Women's Health Issues* 12: 138-49.
- Hornbrook, M. C., A. V. Hurtado, and R. E. Johnson. 1985. "Health Care Episodes: Definition, Measurement and Use." *Medical Care Review* 42: 163-218.

- Hu, L. T., and P. M. Bentler. 1995. "Evaluating Model Fit." In *Structural Equation Modeling Concepts, Issues, and Applications*, edited by R. H. Hoyle, pp. 37-75. Thousand Oaks, CA: Sage.
- Joreskog, K. G., and D. Sorbom. 1984. *Lisrel-VI User's Guide*. 3d Edition. Mooresville, IN: Scientific Software.
- Lee, S. Y., and K. L. Tsui. 1982. "Covariance Structure Analysis in Several Populations." *Psychometrika* 47: 297-308.
- Linder-Pelz, S. 1982. "Toward a Theory of Patient Satisfaction." *Social Science and Medicine* 16: 577-82.
- Linhart, H., and W. Zucchini. 1986. *Model Selection*. New York: Wiley.
- Magill, M., and J. Senf. 1987. "A New Method for Measuring Continuity of Care in Family Practice Residencies." *Journal of Family Practice* 24: 165-8.
- Marshall, G. N., R. D. Hays, and R. M. Rand. 1996. "Health Status and Satisfaction with Health Care: Results from the Medical Outcomes Study." *Journal of Consulting and Clinical Psychology* 64: 380-90.
- Michalos, A. C. 1985. "Multiple Discrepancies Theory (MDT)." *Social Indicators Research* 16: 347-413.
- National Committee for Quality Assurance. 1998. *HEDIS 1999 Volume 3: HEDIS Protocol for Administering CAHPS 2.0H Survey*. Washington, DC: NCQA.
- O'Malley, A. S., C. B. Forrest, and P. G. O'Malley. 2000. "Low-Income Women's Priorities for Primary Care." *Journal of Family Practice* 49: 141-6.
- Peck, B. M., D. A. Asch, S. D. Goold, D. L. Roter, P. A. Ubel, L. M. McIntyre, K. H. Abbott, J. A. Hoff, C. M. Koropchak, and J. A. Tulsky. 2001. "Measuring Patient Expectations: Does the Instrument Affect Satisfaction or Expectations?" *Medical Care* 39: 100-8.
- Pedhazur, E. J. 1982. *Multiple Regression in Behavioral Research*. 2d Edition, pp. 230-2. New York: Holt.
- Rao, J. K., M. Weinberger, L. A. Anderson, and K. Kroenke. 2004. "Predicting Reports of Unmet Expectations among Rheumatology Patients." *Arthritis and Rheumatism* 41: 215-21.
- Schafer, J. L. 1997. *Analysis of Incomplete Multivariate Data*. New York: Chapman & Hall.
- Scholle, S. H., C. S. Weisman, R. T. Anderson, and F. Camacho. 2004. "The Development and Validation of a Woman-Specific Patient Satisfaction Tool: The Primary Care Satisfaction Survey for Women (PCSSW)." *Women's Health Issues* 14: 35-50.
- Sitzia, J., and N. Wood. 1997. "Patient Satisfaction: A Review of Issues and Concepts." *Social Science and Medicine* 45: 1829-43.
- Strasser, S. 1993. "The Patient Satisfaction Process: Moving toward a Comprehensive Model." *Medical Care Review* 50 (2): 219-48.
- Ware, J. E., and R. D. Hays. 1988. "Methods for Measuring Patient Satisfaction with Specific Medical Encounters." *Medical Care* 26: 393-402.
- Weisman, C. S. 1996. "Women's Use of Health Care." In *Women's Health: The Commonwealth Fund Survey*, edited by M. M. Falik and K. S. Collins, pp. 19-48. Baltimore: Johns Hopkins University Press.

- Weisman, C. S., and J. T. Henderson. 2001. "Managed Care and Women's Health: Access, Preventive Services, and Satisfaction." *Women's Health Issues* 11: 201-15.
- West, G. W., J. F. Finch, and P. J. Curran. 1995. "Structural Equation Models with Nonnormal Variables." In *Structural Equation Modeling Concepts, Issues, and Applications*, edited by R. H. Hoyle, pp. 37-75. Thousand Oaks, CA: Sage.
- Wright, S. 1934. "The Method of Path Coefficients." *Annals of Mathematical Statistics* 5: 161-215.