

Psychometric Properties of the Consumer Assessment of Health Plans Study (CAHPS[®]) 2.0 Adult Core Survey

J. Lee Hargraves, Ron D. Hays, and Paul D. Cleary

Objective. To estimate the reliability and validity of survey measures used to evaluate health plans and providers from the consumer's perspective.

Data Sources. Members (166,074) of 306 U.S. health plans obtained from the National CAHPS[®] Benchmarking Database 2.0, a voluntary effort in which sponsors of CAHPS[®] surveys contribute data to a common repository.

Study Design. Members of privately insured health plans serving public and private employers across the United States were surveyed by mail and telephone. Interitem correlations and correlations of items with the composite scores were estimated. Plan-level and internal consistency reliability are estimated. Multivariate associations of composite measures with global ratings are also examined to assess construct validity. Confirmatory factor analysis is used to examine the factor structure of the measure.

Findings. Plan-level reliability of all CAHPS[®] 2.0 reporting composites is high with the given sample sizes. Fewer than 170 responses per plan would achieve plan-level reliability of .70 for the five composites. Two of the composites display high internal consistency (Cronbach's alpha $\geq .75$), while responses to items in the other three composites were not as internally consistent (Cronbach's alpha from .58 to .62). A five-factor model representing the CAHPS[®] 2.0 composites fits the data better than alternative two- and three-factor models.

Conclusion. Two of the five CAHPS[®] 2.0 reporting composites have high internal consistency and plan-level reliability. The other three summary measures were reliable at the plan level and approach acceptable levels of internal consistency. Some of the items that form the CAHPS[®] 2.0 adult core survey, such as the measure of waiting times in the doctor's office, could be improved. The five-dimension model of consumer assessments best fits the data among the privately insured; therefore, consumer reports using CAHPS[®] surveys should provide feedback using five composites.

Key Words. Consumer assessments, CAHPS[®], patient survey, psychometrics, health plan report cards

Systematic collection of consumer feedback is now common in health care. Health plans and providers use survey data for quality assessment and improvement; purchasers use consumer surveys for selecting health plans;

and, consumers and patients are beginning to use survey data to choose health professionals, group practices, and health plans (Buchner and Probst 1999; Finkelstein, Harper, and Rosenthal 1999; Hargraves et al. 1993; Hays et al. 1998; Carman et al. 1999; Hibbard et al. 2000; Knutson et al. 1998; Turnbull and Hembree 1996; Veroff et al. 1998). The Consumer Assessment of Health Plans Study (CAHPS[®]) surveys are currently the most widely used survey instruments that ask consumers about experiences with and evaluations of ambulatory care received from health care professionals and health plans. The CAHPS[®] surveys are being used by state Medicaid programs, employer groups, the Medicare Program, the Federal Employees Health Benefits Program, and a wide range of health plans. In 1999, over 90 million Americans had access to information on health plans derived from CAHPS[®] surveys (Agency for Healthcare Research and Quality 2000; Cleary 1999).

The CAHPS[®] survey reports provide information about health plans and providers. The purpose of the reports is to allow consumers to judge different aspects of health plan performance and assist their selection effort (McGee et al. 1999; Spranca et al. 2000). Because too much information may inhibit consumers' abilities to use reports to aid their decision-making (Hibbard, Slovic, and Jewett 1997), CAHPS[®] 2.0 protocols recommend that individual survey items be combined into five reporting composites with the following characteristics: (1) items within each reporting composite reflect consumer assessments of either health care professionals or health plans in conceptually similar domains and (2) the items within each reporting composite all have the same response options.

Each of the CAHPS[®] questions asks about different aspects of getting medical care; some relate more to activity in the doctor's office and other questions are related characteristics of the health plan. For reporting to consumers, responses to these questions are summarized in composite indexes. These indexes were not designed to be necessarily unidimensional measures of health plan performance. Rather, they were designed to comprise information that would appear to be coherent to consumers. For example, responses to getting needed care and getting care quickly are combined in

This work was supported in part by cooperative agreement HS09205 from the Agency for Healthcare Research and Quality (formerly the Agency for Health Care Policy and Research).

Address correspondence to J. Lee Hargraves, Ph.D., Center for Studying Health System Change, 600 Maryland Ave., Suite 550, Washington, DC 20024. Ron D. Hays, Ph.D., is with the UCLA School of Medicine and RAND, Health Sciences Program. Paul D. Cleary, Ph.D., is with the Department of Health Care Policy, Harvard Medical School, Boston.

separate reporting composites even though those items were highly correlated with each other, because consumers usually view these as separate issues.

The first version of the survey was analyzed to develop reporting composites (Harris-Kojetin et al. 1999; Hays et al. 1999; Marshall et al. 2001). The CAHPS[®] 1.0 survey instrument exhibited excellent plan-level and acceptable internal consistency reliability among privately insured and Medicaid consumers. The reliability and validity of version 2.0 of the survey and reporting composites have not been previously reported.

This paper reports the reliability and validity of the CAHPS[®] 2.0 reporting composites using analysis of data from the National CAHPS[®] Benchmarking Database (NCBD) 2.0, a voluntary effort in which sponsors of CAHPS[®] surveys contribute their data to a common repository. The paper uses data from individuals in health plans serving public and private employers to assess plan-level and internal consistency reliability, and the construct validity of the five composites recommended to summarize consumer experiences with health plans and health professionals. Models containing fewer CAHPS[®] composites have been proposed (Bender and Garfinkel 2001; Zaslavsky et al. 2002). This paper compares alternate factor models to the five-composite model. Fewer composites may reduce the burden among consumers assessing health plan performance; however, we hypothesize that the five-factor model fits the data better than models containing fewer factors. Additionally, we expect that although internal consistency reliability may be lower than desirable for scale construction, plan-level reliability of the CAHPS[®] composite measures is sufficient to make comparisons among health plans. Because CAHPS[®] surveys help consumers to evaluate how plans match up to one another, plan-level reliability is key to the utility of the consumer reports of health plans.

METHODS

Data Source

The data used for these analyses were obtained from the National CAHPS[®] Benchmarking Database (NCBD) 2.0 and represent 166,074 privately insured respondents from 306 health plans. The data represent the results of 13 purchaser and health plan sponsors of surveys that contributed data to the NCBD for 1999. These surveys, fielded using a combination of mail and telephone methods, were conducted between January and June of 1999, with the exception of one sponsored survey of 11 plans completed in 1998.

Response rates ranged from 24 to 57 percent, at the plan level. Although most health plans were health maintenance organizations ($n = 277$), some preferred provider organization, point-of-service, and fee-for-service plans were included in the database ($n = 29$).

Measures

The CAHPS[®] 2.0 questionnaire contains 43 items, of which 19 are core items that are routinely reported to consumers. These 19 items include two global ratings of care that use a 0 to 10 rating scale and 17 items that ask consumers to report about their experiences with health professionals and health plans. The remaining questions in the CAHPS[®] 2.0 survey instrument include two additional global rating items, questions that ask about members' health plan usage, demographic questions, and screening questions to identify questions that may not apply to all survey respondents.

The CAHPS[®] 2.0 survey questions typically are grouped into five composites for public reporting: *Getting Care Quickly*, *Doctors Who Communicate Well*, *Courteous/Helpful Office Staff*, *Getting Needed Care*, and *Health Plan Customer Service*. These five composites also were used with CAHPS[®] 1.0; however, CAHPS[®] items have been modified based on field studies, additional cognitive interviews, and testing with consumers. Several questions used in the first version of the survey were refined or dropped. The CAHPS[®] 2.0 survey also contains additional items. The content of the revised survey is described below. All questions use a 12-month recall period for persons with private insurance.

Getting Care Quickly. The CAHPS[®] 2.0 survey includes four questions about getting help when calling the office, getting appointments for routine care, getting appointments right away for illness or injury, and waiting in the doctor's office longer than 15 minutes. All questions in this composite use the *never to always* response format (i.e., *never, sometimes, usually, always*). The question asking about how often the consumer waits in the doctor's office over 15 minutes is the sole survey item that is reverse coded; the "always" response indicates a worse assessment of the doctor's office.

Doctors Who Communicate Well. Three items assess communication of doctors with patients (listening carefully, explanations, and respect for what is said). The fourth item in this composite asks if the doctor spent enough time with the patient. All of these questions use the *never to always* response task.

Courteous/Helpful Office Staff. Two questions use the *never to always* response task and ask how often the office staff at the doctor's office treated the consumer with courtesy and respect, and how helpful they were.

Getting Needed Care. There are four questions that ask about problems in getting care. Using a *not a problem, a small problem, or a big problem* response scale, the respondent is asked about problems getting a personal doctor or nurse, getting a referral to a specialist, getting care the patient or doctor thought necessary, and delays while waiting for approval from the health plan.

Health Plan Customer Service. The CAHPS[®] 2.0 survey includes three questions about plan customer service that ask about problems with finding and understanding information in the plan's written material, getting help when calling the plan's customer service, and paperwork for the health plan. These three questions use the *not a problem to big problem* response format.

Each of the 17 questions that form the five reporting composites is preceded by a screening question to ensure that only those persons who had the service from the health plan or provider are included. The table located in the Appendix shows CAHPS[®] 2.0 survey items and response options in their respective reporting categories.

The CAHPS[®] survey contains questions that are asked of all consumers (e.g., global rating of the plan) and questions that are asked of subgroups (e.g., getting care for illness or injury). Hence, some data for this analysis is appropriately missing and a small amount of data is inappropriately missing. The total amount of missing data for 22 items used in this analysis ranged from 3 percent for the global rating of the health plan to 66 percent for problems with health plan paperwork (AC37, Appendix). Values for both appropriately and inappropriately missing responses were imputed using a hot-deck strategy data to avoid sample lost due to case-wise deletion in the analysis of composite measures (Rubin 1987; Brick and Kalton 1996). We used the same method for data imputation as applied in an examination of CAHPS[®] version 1.0 reporting composites (Marshall et al. 2001). This imputation method replaces missing data for individual respondents, assuming that they would have responded much like individuals with otherwise comparable ratings of overall health care. Respondents were grouped into quintiles based on the average of their responses to two global rating items (i.e., rating of health care and rating of health plan). Stratifying by these global ratings limits the random drawing of values to groups of individuals who were similar in their overall evaluations of their health plans (Marshall et al. 2001). For each missing item, a randomly selected value was drawn without replacement from among respondents in the same quintile who had answered that item. The randomly selected value was then used in the place of the missing data point for each individual.

The CAHPS[®] analysis methods recommend collapsing responses in all items using the *never to always* response option to obtain a 3-category

distribution, so we combined *never* and *sometimes* into a single category (Agency for Healthcare Policy and Research 1999). All questions using the problem response set (*not a problem, a small problem, a big problem*) were similarly coded from 1 to 3. Hence, all questionnaire items that contribute to reporting composites were coded from 1 to 3, where higher scores indicate more positive assessments of either health professionals or health plans.

Global Rating Items. The survey contains four global rating items, two relate to personal doctors or specialists and two relate to all health care and the health plan. Each global rating asks the consumer to rate their care from 0 to 10, where 0 represents the worst possible care and 10 represents the best possible care.

Data Analysis

We calculated a measure of internal consistency, Cronbach's alpha, to estimate the reliability of the reporting composites. Alpha coefficients greater than .70 are considered indicative of good reliability (Nunnally 1978). The CAHPS[®] 2.0 survey was designed as a tool for measuring health plan performance; therefore, assessing plan-level reliability is a key consideration for users of the survey composites. This measure of reliability assesses for each of the five composites whether the variation within health plans detracts from the variation between plans. In other words, this reliability index represents the ratio of the variance between health plans over the sum of the between-plan variance plus measurement error (Solomon et al. 2002). We assessed plan-level reliability using a one-way analysis of variance with health plans as the between factor: $(MS_{between} - MS_{within})/MS_{between}$. We calculated bivariate correlation coefficients to examine the association between an individual item and its reporting composite, correcting for item overlap with the composite score. We also report the number of respondents needed to achieve acceptable levels of plan-level reliability. We also report the correlations of each reporting composite and items with the global rating items.

We conducted confirmatory factor analysis using the CALIS procedure available in *SAS* for Windows, release 8 (SAS Institute 1990). This procedure uses maximum likelihood estimation to solve simultaneously a series of regression equations that produce an estimated covariance matrix. The distributions of the individual-level items depart from normality. However, maximum likelihood estimation tends to be robust to departures from multivariate normality (Huba and Harlow 1987).

Because of the large sample size, all models are statistically rejectable. We present several practical goodness-of-fit indices including the goodness-of-fit index adjusted for degrees of freedom (AGFI), comparative fit index (Bentler 1990), and the normed fit index (NFI) (Bentler and Bonett 1980). These indexes compare the observed sample covariance matrix against the matrix estimated from the model relative to a null model.

The hypothesized factor models were based on three assumptions. First, we assumed that items contributed information to one, and only one, factor in the model. Second, the factors were allowed to be correlated. Finally, the variance of each factor was fixed at 1.0. We tested three different factor structures. We fit a five-factor model that corresponds to the recommended CAHPS[®] reporting composites. We then constrained the correlation coefficient between two sets of factors to be equal to 1.0. This modification of the factor structure resembles the three-factor model proposed by Bender and Garfinkel (2002). This model combines the CAHPS[®] composites *Getting Care Quickly* and *Getting Needed Care* into a general measure of timely access to care. Two CAHPS[®] composites, *Doctors Who Communicate Well* and *Courteous/Helpful Office Staff*, were combined into a single measure of the quality of provider or staff communications. Last, by setting the correlation between the three factors that represented assessments of health professionals and the correlation between the two factors related to the health plan to 1, we produced a two-factor model that separates those items that may relate more to health plans from those that are more related to providers. This two-factor structure contains one factor with all CAHPS[®] items that use the *never, sometimes, usually, always* response format and the other factor with all CAHPS[®] items that use the *no, small, big problem* response format. Because many users would like to present information to consumers in the simplest possible format, the two-factor model could be used to more succinctly summarize the performance of the health plan and health professionals. Table 2 contains all the CAHPS[®] items and shows the combinations used to produce the three different factor structures.

Both the two- and three-factor models can be viewed as models that are hierarchically related to the five-factor model. Hence, one can calculate a difference in chi-squared statistics with degrees of freedom (DF) equal to the DF in the five-factor model minus the DF in the constrained models. This difference in chi-squared values provides information about the statistical significance of differences between the alternate factor structures.

To examine the associations between reports about experiences in receiving care and the global ratings of personal doctors, specialists, all health

care, and health plans, we estimated four separate regression models in which the global ratings of care were regressed on the five reporting categories. We hypothesized that the composites *Getting Care Quickly*, *Communication with Doctors*, and *Doctor's Office Staff* would be more strongly associated with global ratings of personal doctors or nurses, specialists, and care received from all health professionals than with ratings of health plans. Furthermore, we hypothesized that the *Getting Needed Care* and *Plan Customer Service* composites would be more strongly associated with global ratings of health plans than with global ratings of health professionals.

RESULTS

Empirical Evaluation of CAHPS[®] Reporting Composites

Intercorrelations of Items, Composites, and Ratings. Items in the "getting needed care" composite were modestly associated with the scale score, correcting for item overlap (Table 1). The item-to-scale corrected correlations ranged from .27 to .42. The four items, as well as the composite measure, were more strongly associated with global ratings of all health care than with global ratings of the health plan, personal doctors, or specialists.

The four items in the communication composite were uniformly associated with the scale score, with r 's ranging from .68 to .75. These items were also more strongly correlated with global ratings of all health care than with the other ratings questions. The composite score was more strongly associated with the global rating of all health care ($r = .58$) than with the global ratings of personal doctors, specialists, or health plans (r 's = .41, .26, and .38, respectively).

The two items that comprise the composite assessing doctor's office staff were positively associated with one another ($r = .61$). These two items, as the ones in the previously discussed composites, were moderately associated with the global health care rating and weakly associated with the other three global ratings.

The *Getting Needed Care* composite had similar associations with the global ratings of health care and health plan. Moreover, items in the composite tended to have similar association with the global ratings of health care and health plans (i.e., r 's ranged from .31 to .44). The items related to getting needed care were all moderately associated with the scale score. Item-scale correlations were highest for the item assessing problems getting care that the patient and doctor believed necessary ($r = .48$).

Table 1: Correlations of CAHPS® Adult Survey Items with Reporting Composite and Global Ratings

Adult Survey Composites and Items	Global Ratings of				
	CAHPS® Reporting Composite†	Personal Doctor‡	Specialist‡	All Health Care‡	Health Plan‡
Getting Care Quickly	1.0	.40	.28	.57	.48
AC15 how often got help when calling doctor	.42	.34	.22	.46	.38
AC17 how often got appointment for regular care	.40	.27	.19	.38	.32
AC19 how often got care for illness or injury	.36	.27	.20	.40	.36
AC24 how often waited < 15 minutes	.27	.19	.13	.28	.22
Doctors Who Communicate Well	1.0	.46	.30	.66	.44
AC27 how often doctors listened carefully	.73	.41	.26	.58	.38
AC28 how often doctors explained things	.68	.35	.24	.51	.33
AC29 how often doctors showed respect	.75	.40	.26	.57	.38
AC30 how often doctors spent enough time	.69	.39	.26	.56	.38
Courteous/Helpful Office Staff	1.0	.34	.23	.52	.38
AC25 how often office staff courteous/respectful	.61	.28	.19	.43	.31
AC26 how often office staff was helpful	.61	.33	.23	.50	.37
Getting Needed Care	1.0	.35	.26	.54	.57
AC06 no problem getting personal doctor or nurse	.32	.29	.19	.39	.41
AC10 no problem getting referral to specialists	.40	.24	.18	.36	.38
AC22 no problem getting care thought necessary	.48	.27	.20	.44	.38
AC23 no problem with delays for plan approval	.43	.17	.14	.31	.41
Health Plan Customer Service	1.0	.28	.24	.43	.62
AC33 no problem in finding or understanding written plan information	.30	.20	.16	.30	.40
AC35 no problem getting help from customer service	.35	.20	.18	.32	.48
AC37 no problem with paperwork for plan	.32	.20	.17	.31	.44

Source: Responses (166,074) to CAHPS® surveys among 306 plans in the 1999 National CAHPS® Benchmarking Database. All question numbers refer to the CAHPS® 2.0 Adult Survey.

Notes: †Item to composite correlations were corrected for item overlap with the composite score.

‡Each rating ranged from 0 to 10, where 0 is the worse possible and 10 is the best possible.

All correlation coefficients were significantly different from zero, $p < .0001$.

The three items that form the health plan customer service measure were uniformly and weakly associated with the scale score. These items showed stronger correlations with the global rating of the health plan than with the global ratings of doctors and all health care. As with all of the five composites, the scale score was more strongly associated with the global rating than any of the items that form the composite ($r = .62$).

Confirmatory Factor Analysis. Table 2 presents the factor loadings for five-, three-, and two-factor models. Measures of goodness of fit are shown below the table. In the five-factor model, the four items in the *Getting Care Quickly* composite loaded modestly on the first factor. The lowest factor loading in this composite was for the item assessing how often the consumer waited more than 15 minutes in the doctor's office to see the person he or she came to see. This item was reverse coded so that a positive loading indicates that the respondent was more likely to say that the wait was less than 15 minutes. The item related to problems getting a personal doctor or nurse contributed less to the *Getting Needed Care* composite (factor 2) than the other three items; the item assessing problems getting necessary care contributed most to this composite.

Each item in the *Doctor's Who Communicate Well* composite was strongly associated with the third factor. Similarly, the two items in the *Helpful/Courteous Office Staff* composite were strongly associated with factor 4. The items in the *Plan Customer Service* composite (factor 5) had factor loadings that were modest in magnitude (from .48 to .55).

The five-factor model fit the data well and had slightly better fit than both the two- and three-factor models. For example, the goodness-of-fit index adjusted for degrees of freedom (AGFI) for the five-factor model was .99, whereas the AGFI for the three-factor model was .96, and the two-factor model was .95. The three-factor model was significantly different from the five-factor model (χ^2 difference = 37,935; $df = 2$; $p < .001$). Similarly, the two-factor model was markedly different than the five-factor model (χ^2 difference = 51,628; $df = 4$; $p < .001$).

In the three-factor model, shown in Table 2, all factor loadings for the access to care factor are between .35 and .62 in magnitudes similar to the two factors in the CAHPS[®] composites. The pattern of the factor loadings for the second factor appears to maintain two groups that match the original CAHPS[®] composites. For example, the loadings for the items related to office staff (i.e., courteous/respectful and helpful) were both less than .7, while the loadings for the four items related to doctor's communication were all greater than .7. The third factor in this model was identical to the one presented in the five-factor model.

Table 2: Factor Loadings from Confirmatory Factor Analysis (CFA) of CAHPS® 2.0 Adult Survey Report Items

CAHPS® Survey Item	Factors from Separate CFA Models									
	Five					Three			Two	
	1	2	3	4	5	1	2	3	1	2
AC15 how often got help when calling doctor	.63					.56				.52
AC17 how often got appointment for regular care	.53					.47				.44
AC19 how often got care for illness or injury	.51					.47				.42
AC24 how often waited < 15 minutes	.39					.35				.35
AC06 no problem getting personal doctor/nurse		.47				.47				.47
AC10 no problem getting referral to specialists		.52				.49				.51
AC22 no problem getting necessary care		.66				.62				.62
AC23 no problem with delays for approval		.57				.52				.56
AC27 how often doctors listened carefully			.82				.81			.80
AC28 how often doctors explained things			.74				.73			.71
AC29 how often doctors showed respect			.82				.81			.79
AC30 how often doctors spent enough time			.75				.75			.74
AC25 how often office staff was courteous and respectful				.72			.57			.60
AC26 how often office staff was helpful				.84			.66			.68
AC33 no problem in finding or understanding written plan information					.48			.48		.41
AC35 no problem getting help from customer service					.55			.55		.45
AC37 no problem with paperwork for plan					.50			.50		.42

Note: All factor loadings were significantly greater than zero, $p < .0001$.

Five-factor model: AGFI = .99; CFI = .98; NFI = .98; RMSEA = .03; $\chi^2 = 12,854.21$; $df = 109$.
 Three-factor model: AGFI = .96; CFI = .94; NFI = .94; RMSEA = .05; $\chi^2 = 50,789.26$; $df = 111$.

Two-factor model: AGFI = .95; CFI = .92; NFI = .92; RMSEA = .06; $\chi^2 = 64,482.26$; $df = 113$.

AGFI = Goodness-of-Fit Index Adjusted for Degrees of Freedom; CFI = Comparative Fit Index; NFI = Normed Fit Index; and RMSEA = Root Mean Squared Error of Approximation.

Table 3: Estimated Correlations among CAHPS[®] Dimensions

Factor	Five-Factor Model			
	1	2	3	4
1. Getting Care Quickly				
2. Doctors Who Communicate Well	.74			
3. Helpful/Courteous Office Staff	.75	.74		
4. Getting Needed Care	.75	.58	.53	
5. Health Plan Customer Service	.64	.49	.45	.72

Source: Responses from surveys among 306 plans in the NCBD, CAHPS[®] 2.0 Adult Survey.

The factor loadings for the two-factor model (see Table 2) indicate items that form the “doctors who communicate” and “helpful/courteous office staff” composites contribute more to factor 1 than the four items from the getting care quickly composite. The lowest factor loading in this two-factor model was for the item assessing waits in the doctor office more than 15 minutes. The two items related to problems getting necessary care and delays for plan approval contribute more to the second factor than the other seven items. The lowest two-factor loadings for factor 2 were for the item related to plan paperwork (AC37) and for plan written information (AC33).

Table 3 shows the estimated interfactor correlations for the five-factor model. The first factor representing getting care quickly was more positively associated with the composites related to getting care, communicating with doctors, and doctor’s office staff than with the plan customer service composite. The “getting needed care” and “plan customer services” factors were positively associated ($r = .72$). The estimated correlation between the health plan and provider factors in the two-factor model was .66.

Reliability and Validity of CAHPS[®] Composites and Ratings

Reliability of Report Composites. Table 4 contains measures of plan-level reliability and internal consistency. All of the CAHPS[®] composites have high plan-level reliability. The number of responses needed to achieve plan-level reliability of greater than .70 was highest for the “doctors who communicate well” and “helpful/courteous office staff composites” (i.e., more than 140 responses). Fewer than 90 responses would be needed for the other three composites to achieve adequate plan-level reliability.

Table 4: Reliability Estimates for CAHPS[®] Reporting Composites and Global Rating Items

<i>Adult Survey Composites and Ratings</i>	<i>Plan-Level Reliability (Mean Sample Size = 543)</i>	<i>Number of Responses to Obtain Adequate Plan-Level Reliability[†]</i>	<i>Internal Consistency[‡] (N = 166,074)</i>
Getting Care Quickly	.94	84	.58
Doctors Who Communicate	.88	169	.86
Courteous/Helpful Office Staff	.90	141	.75
Getting Needed Care	.95	64	.62
Customer Service	.94	83	.51
Global Ratings of			
Personal Doctor or Nurse	.88	173	N/A
Specialist	.82	287	N/A
Quality of Health Care	.93	98	N/A
Health Plan	.96	49	N/A

Source: Responses from surveys among 306 plans in the NCBDB, CAHPS[®] 2.0 Adult Survey; mean sample size per plan was 543, median was 514.

Notes: [†]Using the Spearman-Brown prophecy formula to achieve reliability greater than .7.

[‡]To include all respondents, alpha coefficients were calculated using imputed data. Cronbach's alpha coefficients calculated using listwise or pairwise deletion of cases without imputation were all greater than .70. Internal consistency coefficients were .73, .89, .80, .71, and .72, respectively.

The two composites showing the highest internal consistency were related to how well doctors communicate and the helpfulness and courtesy of the doctor's office staff (Cronbach's alpha of .86 and .75, respectively). The remaining three composites had internal consistency coefficients less than .70, with the plan customer service composite having the lowest internal consistency.

Reliability of Global Ratings. Plan level reliability was high for the four global ratings (Table 4). Global ratings of overall care and health plan were more reliable than the ratings of personal doctors and specialists. The two ratings of doctors require more responses per plan to achieve adequate plan-level reliability than the ratings of all health care.

Association between Composites and Ratings. Table 5 displays standardized regression coefficients from linear models in which the four global rating items were regressed on the CAHPS[®] composites. The smallest amount of explained variance in the dependent variables was for the global rating of specialty care ($R^2 = .13$) and the largest for the rating of health care overall ($R^2 = .56$). The *Doctors Who Communicate Well* composite was the strongest predictor of ratings of specialists ($\beta = .17$), personal doctors ($\beta = .29$), and

Table 5: CAHPS[®] Global Rating Items Regressed on Reporting Composites

<i>Adult Survey Composites</i>	<i>Number of Items</i>	<i>Standardized Regression Coefficients[†]</i>			
		<i>Personal Doctor or Nurse</i>	<i>Specialist</i>	<i>All Health Care</i>	<i>Health Plan</i>
Getting Care Quickly	4	.14	.10	.18	.15
Doctors Who Communicate	4	.29	.17	.38	.08
Courteous/Helpful Office Staff	2	.03	.02	.09	.03
Getting Needed Care	4	.12	.09	.20	.28
Customer Service	3	.07	.12	.13	.42
<i>R²</i>		.26	.13	.56	.54

Source: Responses from surveys among 306 plans in the NCBH, CAHPS[®] 2.0 Adult Survey.

Notes: [†]Each rating ranged from 0 to 10, where 0 is the worst possible and 10 is the best possible. All standardized regression coefficients were significantly different from zero, $p < .0001$, $n = 166,074$

health care ($\beta = .38$). The two composites assessing ability to get care and get it quickly were weakly associated with global ratings of personal doctors, adjusting for the other composites.

The best predictor of the health plan rating was the plan customer service composite ($\beta = .42$). The plan customer service composite was weakly and positively associated with the rating of all health care. Getting care quickly, courteous office staff, and getting needed care, were associated with the global ratings of all health care and the health plan in the same direction and at about the same magnitude. For example, the standardized coefficient for the “getting needed care” composite was .20 in the all health care model and .28 in the health plan model. Hence, these three composites provide information related to all health care as well as the health plan rating.

DISCUSSION

The CAHPS[®] composites were created to summarize complex information for users of reports about health plan performance, mostly consumers. They were not designed to be internally consistent scales. Nonetheless, the scales tend to be internally consistent. Two of the reporting composites have very high internal consistency and three were lower. However, all five reporting composites displayed impressive plan-level reliability.

The item that most appears to need refinement is the one asking about wait times at the doctor’s office. This item was weakly correlated with the

“getting care quickly” scale score and the global rating of care. This may be because it is the only negatively worded CAHPS[®] question. Deleting this item had almost no impact on the internal consistency of the composite (i.e., a Cronbach’s alpha of .58 would be observed in a three-item index). Deleting this item also had no effect on the plan-level reliability.

Each of the five CAHPS[®] reporting composites was positively associated with global ratings of health care and health plans. Furthermore, the correlations among these five reporting composites were moderate to high. It is likely that consumers use experiences with their health professionals when rating their plans or experiences with plans may influence ratings of professionals. Nevertheless, communication of doctors and health professionals was the strongest correlate of consumer ratings of health care and the performance of health plan customer service was the best correlate of consumer ratings of the health plan.

To reduce the amount of information presented to consumers, CAHPS[®] survey sponsors (e.g., health plans, employers, or other purchasers) could create two composite measures. The two- and three-factor models displayed reasonably good fit in the confirmatory factor analysis. However, the five-factor model representing the CAHPS[®] reporting composites displayed better fit to the data. For example, factor 1 in the two-factor model had six items that loaded greater than .60 and four items that loaded less than .55. Thus, the first factor in this simplified model appeared to represent two domains. Although users may wish to present the results of two reporting composites (i.e., health plans and health professionals), some information about health professional’s performance will be masked to consumers. To provide consumers with information about distinct domains of health plan performance, users of CAHPS[®] surveys should consider presenting the five reporting composites we examined among privately insured individuals. Considerations also should be given to the extent to which the different aggregation schemes adequately represent variability across plans in average CAHPS[®] scores (Zaslavsky et al. 2000).

Other examinations of the CAHPS[®] surveys propose presenting fewer reporting composites with other measures created directly from the survey (Bender and Garfinkel 2001) or with information on clinical quality (Zaslavsky et al. 2002). In some circumstances, consumers might be overwhelmed with five composite measures from the survey presented with performance measures from expanded CAHPS[®] surveys or other data sources (e.g., Medicare prevention ratings or HEDIS measures). Hence, fewer CAHPS[®] composite measures of health plan performance to consumers might be a

necessity. Developers of consumer reports of health plan performance must balance the desire for comprehensive information about distinct characteristics of health plans reported in narrow categories with the necessity of reducing cognitive burden among readers.

APPENDIX

CAHPS[®] 2.0 Adult Survey Composites (Response Options) and Items

Getting Care Quickly (never, sometimes, usually, always)

- AC15 In the last 12 months, when you called during regular office hours, how often did you get the help or advice you needed?
- AC17 In the last 12 months, how often did you get an appointment for regular or routine care as soon as you wanted?
- AC19 In the last 12 months, when you needed care right away for an illness or injury, how often did you get care as soon as you wanted?
- AC24 In the last 12 months, how often did you wait in the doctor's office or clinic more than 15 minutes past your appointment time to see the person you went to see?

Doctors Who Communicate Well (never, sometimes, usually, always)

- AC27 In the last 12 months, how often did doctors or other health professionals listen carefully to you?
- AC28 In the last 12 months, how often did doctors or other health professionals explain things in a way you could understand?
- AC29 In the last 12 months, how often did doctors or other health professionals show respect for what you had to say?
- AC30 In the last 12 months, how often did doctors or other health professionals spend enough time with you?

Courteous and Helpful Office Staff (never, sometimes, usually, always)

- AC25 In the last 12 months, how often did office staff at a doctor's office or clinic treat you with courtesy and respect?
- AC26 In the last 12 months, how often were office staff at a doctor's office or clinic as helpful as you thought they should be?

Getting Needed Care (a big problem, a small problem, not a problem)

- AC6 With the choices your health plan gives you, how much of a problem, if any, was it to get a personal doctor or nurse you are happy with?

- AC10** In the last 12 months, how much of a problem, if any, was it to get a referral to a specialist that you wanted to see?
- AC22** In the last 12 months, how much of a problem, if any, was it to get the care you or your doctor believed necessary?
- AC23** In the last 12 months, how much of a problem, if any, were delays in health care while you waited for approval from your health plan?
- Health Plan Customer Service** (a big problem, a small problem, no problem)
- AC33** In the last 12 months, how much of a problem, if any, was it to find or understand information in the written materials?
- AC35** In the last 12 months, how much of a problem, if any, was it to get the help you needed when you called your health plan's customer service?
- AC37** In the last 12 months, how much of a problem, if any, did you have with paperwork for your health plan?
-

Source: CAHPS® 2.0 Adult Survey

REFERENCES

- Agency for Healthcare Policy and Research. 1999. *CAHPS 2.0 Survey and Reporting Kit*. Rockville, MD: Agency for Healthcare Policy and Research.
- Agency for Healthcare Research and Quality. 2000. *From the Pipeline of Health Services Research-CAHPS: The Story of the Consumer Assessment of Health Plans Study. Translating Research into Practice Fact Sheet*. AHRQ publication no. 00-P014. Rockville, MD: Agency for Healthcare Research and Quality.
- Bender, R. H., and S. A. Garfinkel. 2001. "Differences in the Structure of CAHPS Measures among the Medicare Fee-for-Service, Medicare Managed Care, and Privately Insured Populations." *Health Services Research* 36 (3): 489-508.
- Bentler, P. M. 1990. "Comparative Fit Indexes in Structural Models." *Psychological Bulletin* 107 (2): 238-46.
- Bentler, P. M., and D. G. Bonett. 1980. "Significance Tests and Goodness-of-Fit in the Analysis of Covariance Structures." *Psychological Bulletin* 88 (3): 588-606.
- Brick, J. M., and G. Kalton. 1996. "Handling Missing Data in Survey Research." *Statistical Methods in Medical Research* 5 (3): 215-38.
- Buchner, D. A., and L. Y. Probst. 1999. "The Member Satisfaction Survey as a Measure of Health Plan Accountability." *Quality Management in Health Care* 7 (2): 41-9.
- Carman, K. L., P. F. Short, D. O. Farley, J. A. Schnaier, D. B. Elliott, and P. M. Gallagher. 1999. "Epilogue: Early Lessons from CAHPS Demonstrations and Evaluations. Consumer Assessment of Health Plans Study." *Medical Care* 37 (3, Supplement): MS97-105.

- Cleary, P. D. 1999. "The Increasing Importance of Patient Surveys." *Quality in Health Care* 8 (4): 212.
- Finkelstein, B. S., D. L. Harper, and G. E. Rosenthal. 1999. "Patient Assessments of Hospital Maternity Care: A Useful Tool for Consumers?" *Health Services Research* 34 (2): 623-4.
- Hargraves, J. L., R. H. Palmer, J. Zapka, D. Nerenz, H. Frazier, E. J. Orav, C. Warner, J. Ingard, and R. Neisuler. 1993. "Using Patient Reports to Measure Health Care System Performance." *Clinical Performance and Quality Health Care* 1 (4): 208-13.
- Harris-Kojetin, L. D., F. J. Fowler Jr., J. A. Brown, J. A. Schnaier, and S. F. Sweeny. 1999. "The Use of Cognitive Testing to Develop and Evaluate CAHPS 1.0 Core Survey Items. Consumer Assessment of Health Plans Study." *Medical Care* 37 (3, Supplement): MS10-21.
- Hays, R. D., J. A. Brown, W. J. Dixon, and R. H. Brooks. 1998. "Satisfaction with Health Care Provided by 48 Physician Groups." *Archives of Internal Medicine* 158 (7): 785-90.
- Hays, R. D., J. A. Shaul, V. S. Williams, J. S. Lubalin, L. D. Harris-Kojetin, S. F. Sweeny, and P. D. Cleary. 1999. "Psychometric Properties of the CAHPS 1.0 Survey Measures. Consumer Assessment of Health Plans Study." *Medical Care* 37 (3, Supplement): MS22-31.
- Hibbard, J. H., L. Harris-Kojetin, P. Mullin, J. Lubalin, and S. Garfinkel. 2000. "Increasing the Impact of Health Plan Report Cards by Addressing Consumers' Concerns." *Health Affairs* 19 (5): 138-43.
- Hibbard, J. H., P. Slovic, and J. J. Jewett. 1997. "Informing Consumer Decisions in Health Care: Implications from Decision-making Research." *Milbank Quarterly* 75 (3): 395-414.
- Huba, G. J., and L. L. Harlow. 1987. "Robust Structural Equation Models: Implications for Developmental Psychology." *Child Development* 58 (1): 147-66.
- Knutson, D. J., E. A. Kind, J. B. Fowles, and S. Adlis. 1998. "Impact of Report Cards on Employees: A Natural Experiment." *Health Care Financing Review* 20 (1): 5-27.
- McGee, J., D. E. Kanouse, S. Sofaer, J. L. Hargraves, E. Hoy, and S. Kleimann. 1999. "Making Survey Results Easy to Report to Consumers: How Reporting Needs Guided Survey Design in CAHPS. Consumer Assessment of Health Plans Study." *Medical Care* 37 (3, Supplement): MS32-4.
- Marshall, G. N., L. S. Morales, M. Elliott, K. Spritzer, and R. D. Hays. 2001. "Confirmatory Factor Analysis of the Consumer Assessment of Health Plans Study (CAHPS) 1.0 Core Survey." *Psychological Assessment* 13 (2): 216-29.
- Nunnally, J. 1978. *Psychometric Theory*. New York: McGraw-Hill.
- Rubin, D. B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.
- SAS Institute. 1990. *SAS Technical Report P-200, SAS/STAT Software: CALIS and LOGISTIC Procedures, Release 6.04*. Cary, NC: SAS Institute.
- Solomon, L. S., A. M. Zaslavsky, B. E. Landon, and P. D. Cleary. 2002. "Variation in Patient-Reported Quality among Health Care Organizations." *Health Care Financing Review* 23 (4): 85-10.

- Spranca, M., D. E. Kanouse, M. Elliott, P. F. Short, D. O. Farley, and R. D. Hays. 2000. "Do Consumer Reports of Health Plan Quality Affect Health Plan Selection?" *Health Services Research* 35 (5, pt 1): 933-47.
- Turnbull, J. E., and W. E. Hembree. 1996. "Consumer Information, Patient Satisfaction Surveys, and Public Reports." *American Journal of Medical Quality* 11 (1): S42-5.
- Veroff, D. R., P. M. Gallagher, V. Wilson, M. Uyeda, J. Merselis, E. Guadagnoli, S. Edgman-Levitan, A. Zaslavsky, S. Kleimann, and P. D. Cleary. 1998. "Effective Reports for Healthcare Quality Data: Lessons from a CAHPS Demonstration in Washington State." *International Journal of Quality Health Care* 10 (6): 555-6.
- Zaslavsky, A. M., N. D. Beaulieu, B. E. Landon, and P. D. Cleary. 2000. "Dimensions of Consumer-Assessed Quality of Medicare Managed-Care Health Plans." *Medical Care* 38 (2): 162-74.
- Zaslavsky, A. M., J. A. Shaul, L. B. Zaborski, M. J. Cioffi, and P. D. Cleary. 2002. "Combining Health Plan Performance Indicators into Simpler Composite Measures." *Health Care Financing Review* 23 (4): 101-15.

