



Published in final edited form as:

*Med Care*. 2012 November ; 50(Suppl): S28–S34. doi:10.1097/MLR.0b013e31826cbc0d.

## Psychometric Properties of the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) Clinician and Group Adult Visit Survey

**Naomi Dyer, Ph.D.,**

Westat, 1600 Research Blvd, Rockville, MD 20850, 301-610-8842

**Joann S. Sorra, Ph.D.,**

Westat, 1600 Research Blvd, Rockville, MD 20850, 301-294-3933, 301-315-5912 (Fax)

**Scott A. Smith, Ph.D.,**

Westat, 1600 Research Blvd, Rockville, MD 20850, 301-251-8288, 301-315-5912 (Fax)

**Paul Cleary, Ph.D.,** and

Yale University, 60 College Street, P.O. Box 208034, New Haven, CT 0650-8034, 203-785-2867, 203-785-6103 (Fax)

**Ron Hays, Ph.D.**

RAND, Santa Monica, CA, and UCLA Department of Medicine, 911 Broxton Avenue, Room 110, Los Angeles, CA 90024-1736, 310-794-2294, 310-794-0732 (Fax)

Naomi Dyer: NaomiDyer@westat.com; Joann S. Sorra: JoannSorra@westat.com; Scott A. Smith: ScottSmith@westat.com; Paul Cleary: Paul.Cleary@yale.edu; Ron Hays: drhays@ucla.edu

### Abstract

**Background**—The Consumer Assessment of Healthcare Providers and Systems (CAHPSR) Clinician and Group Adult Visit Survey enables patients to report their experiences with outpatient medical offices.

**Objective**—To evaluate the factor structure and reliability of the CAHPS Clinician and Group (CG-CAHPS) Adult Visit Survey.

**Data Source**—Data from 21,318 patients receiving care in 450 clinical practice sites collected from March 2010 to December 2010 were analyzed from the CG-CAHPS Database.

**Research Design and Participants**—Individual level and multilevel confirmatory factor analyses were used to examine CAHPS survey responses at the patient and practice site levels. We also estimated internal consistency reliability and practice site level reliability. Correlations among multi-item composites and correlations between the composites and two global rating items were examined.

**Measures**—Scores on CG-CAHPS composites assessing Access to Care, Doctor Communication, Courteous/Helpful Staff, and two global ratings of whether one would Recommend their Doctor, and an Overall Doctor Rating.

**Results**—Analyses provide support for the hypothesized three-factor model assessing Access to Care, Doctor Communication, and Courteous/Helpful Staff. In addition, the internal consistency reliabilities were 0.77 or higher and practice site level reliabilities for sites with more than four clinicians were 0.75 or higher. All composites were positively and significantly correlated with the

two global rating items, with Doctor Communication having the strongest relationship with the global ratings.

**Conclusions**—The CG-CAHPS Adult Visit Survey has acceptable psychometric properties at the individual level and practice site level. The analyses suggest that the survey items are measuring their intended concepts and yield reliable information.

### Keywords

CAHPS®; consumer; survey; patient experience with care; patient satisfaction; psychometrics; multilevel; confirmatory factor analysis

---

The Consumer Assessment of Healthcare Providers and Systems (CAHPS®) surveys were developed to elicit reports from consumers about their health care experiences. The surveys cover topics such as the communication skills of providers, helpfulness of staff, and access to care, which are important to consumers and for which they are the best source for this information. The surveys and accompanying tools can be used by providers, healthcare organizations, government agencies, and researchers to assess and improve patient-centered care. Establishing the psychometric properties of CAHPS surveys is an integral step toward enabling valid comparisons on patient experience across organizations and over time.<sup>1-5</sup>

The CAHPS Clinician and Group Survey (CG-CAHPS) was developed to assess patient experiences with ambulatory care. There are three versions of CG-CAHPS: 1) a 12-Month Survey that asks patients to report on their experiences over the last 12 months, 2) an expanded 12-Month Survey that includes items to assess aspects of the Patient-Centered Medical Home (PCMH), and 3) a Visit Survey that primarily focuses on experiences during a single visit. The Visit Survey includes questions about doctor communication and office staff interactions at the patient's most recent visit, and questions about the patient's access to care with their doctor over the last 12 months. The survey also elicits an overall rating of the doctor from patients and asks about their willingness to recommend their doctor to family and friends. The Visit Survey was designed to collect feedback about a specific patient visit that providers can use for monitoring and improving care.

In this paper we evaluate the hypothesized factor structure and reliability of the CG-CAHPS Adult Visit Survey using data submitted to the CG-CAHPS Database.

## METHODS

### Measures

The CG-CAHPS Adult Visit Survey contains 42 items, of which 13 are used to create three composites which assess Access to Care (five items), Doctor Communication (six items), and Courteous/Helpful Staff (two items). The survey also includes two questions that ask respondents (1) to rate their doctor, and (2) report if they would recommend the doctor's office to family and friends. In addition, respondents are asked about their overall health, age, gender, and education.

**Access to Care Composite:** The five Access to Care items ask patients about their ability to get an appointment for urgent care as soon as needed, get an appointment for a check-up or routine care as soon as needed, get an answer to a phone question during regular office hours on the same day, get an answer to a phone question after hours as soon as needed, and if the wait time to be seen was within 15 minutes of appointment time. All questions in this composite have a reference period of 12 months and use a four-point response scale (1 = *Never*, 2 = *Sometimes*, 3 = *Usually*, 4 = *Always*). The Access to Care composite uses a 12-

month reference period unlike the other items on the Visit Survey which ask about the most recent visit. In field testing, results showed that the Access items using a visit-based reference period did not achieve an acceptable level of reliability. As a result, the Access items were changed back to the 12-month reference period, leaving all other items visit-specific.

**Doctor Communication Composite:** The six Doctor Communication items ask whether the doctor explained things clearly, listened carefully, gave easy to understand instructions, knew important medical history about the patient, showed respect, and spent enough time with the patient. These questions reference the most recent visit and use a three-point response scale (1 = *Yes, definitely*; 2 = *Yes, somewhat*; 3 = *No*). The items in this composite were recoded such that higher scores equal more positive responses (e.g., *Yes, definitely* was recoded to 3; *No* was recoded to 1).

**Courteous/Helpful Staff Composite:** The two Staffing items ask whether clerks and receptionists were helpful, and if they treated the patient with courtesy and respect. These questions reference the most recent visit and use a three-point response scale (1 = *Yes, definitely*; 2 = *Yes, somewhat*; 3 = *No*). The items in this composite were recoded such that higher scores equal more positive responses (e.g., *Yes, definitely* was recoded to 3; *No* was recoded to 1).

**Overall Doctor Rating:** This question asks the patient to rate the doctor on a scale from 0 to 10, with 0 representing the worst doctor possible and 10 representing the best doctor possible.

**Recommend Doctor Rating:** This question asks whether the patient would recommend the doctor's office to family and friends and uses a three-point response scale (1 = *Yes, definitely*; 2 = *Yes, somewhat*; 3 = *No*). This item was recoded such that higher scores equal more positive responses (e.g., *Yes, definitely* was recoded to 3; *No* was recoded to 1).

### Analysis Dataset

The data was from the CG-CAHPS Database, consisting of 103,442 respondents from 469 practice sites. The Visit Survey includes a number of screener questions that require a "yes" response before responding to a subsequent question. For one of these questions, a majority of respondents (93%) had not phoned their doctor after regular office hours and therefore were instructed to skip the Access to Care item Q12: "When you phoned this doctor's office after regular office hours, how often did you get an answer to your medical question as soon as you needed it?" Because there was such a high percentage of valid skips for this item, it was dropped from further analyses. The remaining Access to Care composite items had responses from between 46% to 98% of the respondents. The two Courteous/Helpful Staff items and five of the six Doctor Communication items were answered by 99%. The Doctor Communication item (Q21) about receiving easy to understand health care instructions was answered by 84% of respondents.

To run a three-factor psychometric model with items loading onto their associated composites (Access, Doctor Communication and Courteous/Helpful Staff), we included only non-missing data for the items that make up the three CG-CAHPS composites. The final analysis dataset therefore consisted of 21,318 responses from 450 practice sites.

### Sample

The data used for these analyses came from health systems, medical offices, and survey vendors who voluntarily submitted CG-CAHPS survey data collected from March 2010 to

December 2010 to the CAHPS Database. All of the 450 practice sites included in the analysis data set administered mail surveys. Most of the practice sites specialized in Family Practice and/or Internal Medicine (89%). Over two-thirds of the practice sites were owned by a hospital or integrated delivery system (69%). Most respondents were female (67%) and a majority were 45 years or older (81%).

## Analyses

Descriptive statistics for the survey items and Spearman rank-order correlations with their associated composites and the global rating items were computed. In addition, we performed confirmatory factor analyses using Mplus Version 6.12, as described below. Finally, we estimated internal consistency reliability and physician group-level reliability (see below).

**Individual-Level Confirmatory Factor Analysis**—We conducted individual-level confirmatory factor analysis on the proposed three-factor model, with maximum likelihood estimation, at first ignoring the nesting of respondents within practice sites. To assess the appropriateness of the resulting structure, we examined factor loadings with the criterion that they should be 0.40 or greater.<sup>6</sup> We present standard overall model fit statistics: the chi-square, comparative fit index (CFI), the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR).

Given the large sample size of our data set, we primarily relied on the CFI, RMSEA, and SRMR as indices of model fit because the chi-square is influenced by sample size such that the larger the sample size the more likely it is that the chi-square will be significant (which indicates lack of model fit).<sup>7, 8</sup> The CFI compares the existing model fit with a null model that assumes the items in the model are uncorrelated. The factor structure is determined to adequately fit the data if the CFI is at least 0.95.<sup>9</sup> The RMSEA examines the residuals of the model; an RMSEA of 0.06 or less is indicative of good fit.<sup>9</sup> The SRMR is the standardized difference between the observed and predicted covariances from the model. A value of zero for the SRMR indicates perfect fit, but a value less than 0.08 is considered good fit.<sup>9</sup>

**Multilevel Analyses**—When respondent data are nested within practice sites, multilevel modeling is generally more appropriate because it accounts for the nested nature of the data. We performed a number of steps in association with the multilevel analyses.

**Intraclass Correlations (ICCs) and Design Effects:** First, we examined intraclass correlations (ICCs) and design effects to determine if the data were truly nested to determine whether multilevel analyses would be necessary.<sup>10</sup> ICCs above 0.05 indicate that the multilevel structure of the data needs to be taken into consideration; ICCs less than 0.05 signify that the consequences of not using multilevel analyses are minimal.<sup>11</sup> We also examined design effects, as ICCs are affected when there are few groups comprised of many individuals or many groups comprised of few individuals, as is the case for our data set. Design effects take into consideration the group sample size (Design Effect = 1 + [Average within group sample size – 1] \* ICC). A design effect of 2.0 or more implies that group membership is associated with responses of the individuals and therefore multilevel modeling should be conducted to account for the multilevel nature of the data.<sup>12</sup>

**Multilevel Confirmatory Factor Analyses (MCFA):** Similar to the individual level confirmatory factor analyses, a three-factor model was examined, taking into consideration the nested nature of the data. We evaluated the item factor loadings with the same rule as the individual level confirmatory factor analyses – that factor loadings should be 0.40 or greater. With multilevel models, two sets of factor loadings are provided: between practice sites and within practice sites, which coincide with the nested nature of the data. The between factor

loadings are based on the between practice site covariance matrix while the within factor loadings use the within or respondent-level covariance matrix. We again present overall model fit indices using standard fit statistics: the chi-square, CFI, RMSEA, and SRMR, with the same criteria as at the individual level.

**Reliability**—Cronbach's coefficient alpha, an estimate of reliability, was calculated for each composite to assess the extent to which respondents consistently answered the items, with a reliability of at least 0.70 considered acceptable for group comparisons.<sup>13</sup>

We examined practice site reliability by practice site size (i.e., the number of clinicians per site) because practices of different sizes need different numbers of patient surveys to reach acceptable levels of reliability on the measures. We calculated practice site reliability using the following formula:

$$\text{Reliability}_g = \frac{\Sigma_B}{\Sigma_B + \frac{\Sigma_W}{N_g}}$$

Where  $\Sigma_B$  refers to the between-group variance;  $\Sigma_W$  refers to the within-group variance, and  $N_g$  is the sample size for practice site  $g$ .<sup>14</sup>

Average reliability estimates were calculated for the three composites and two global rating items for six practice size categories: (1) 1 clinician; (2) 2–3 clinicians; (3) 4–9 clinicians; (4) 10–13 clinicians; (5) 14–19 clinicians, and (6) 20 or more clinicians. A variety of different size categories were considered and other splits are possible but this set of categories was chosen based on variance in reliability and patient sample sizes available in our data set. Similar to internal consistency reliability, values of at least 0.70 are considered acceptable for practice site comparisons.<sup>13</sup>

**Correlations among Composites and Global Ratings**—Relationships among the composites and global ratings at the individual and practice site levels were also examined using Spearman rank-order correlations. While the composites should be correlated since they all measure aspects of patient experience, very high intercorrelations indicate that the composites may not be unique enough to be considered separate measures. In general, composite intercorrelations should be less than 0.80 for the composites to be considered unique.<sup>15</sup> We hypothesized that the composites would be positively related to the global rating items.

## RESULTS

**Descriptive Statistics**—The means, standard deviations, top box scores, and correlations for the survey items are provided in Table 1. Consistent with other patient experience data, CG-CAHPS ratings of care tend to be very positive (negatively skewed)—that is, consumers tend to report positive experiences with health care in the U.S.<sup>16</sup>

The item-to-composite correlations (corrected for item overlap with the composite total) ranged from 0.40 (Q13 with Access to Care) to 0.71 (Q28 and Q29 with Courteous/Helpful Staff). The correlations between the composite items and the global rating items ranged from 0.18 (Q29. Courteous/Helpful Staff with Overall Doctor Rating) to 0.53 (Q19. Doctor Communication with Recommend Doctor).

**Individual Level Confirmatory Factor Analysis**—Table 2 shows that all items within the composites had factor loadings above the 0.40 criteria, with an average loading of 0.68 for Access to Care, 0.76 for Doctor Communication, and 0.86 for Courteous/Helpful Staff. The overall model fit indices are shown in Table 3. As expected, the chi-square test was statistically significant ( $p < 0.01$ ) given the large sample size. The CFI was 0.97, above the 0.95 criterion for good model fit. The RMSEA was 0.05, below the 0.06 criterion, indicating good model fit. The SRMR was 0.04, below the 0.08 criterion, again signifying good model fit. Overall, the individual level factor analysis results provided initial support for the three composites and justification for aggregating the items into their associated composites.

### Multilevel Factor Analyses

**ICCs and Design Effects:** As shown in Table 2, the item ICCs for Access to Care were all above the 0.05 criterion; with an average of 0.08, ranging from 0.07 to 0.11. This finding indicates that between 7% and 11% of the variance may be attributed to practice site membership and establishes the need for multilevel analyses. For Doctor Communication and Courteous/Helpful Staff, all the item ICC values were at or below the 0.05 criterion indicating very little variability across practice sites (average of 0.02, ranging from 0.01 to 0.05). However, when examining design effects, both Courteous/Helpful Staff items and one of the Doctor Communication items had values exceeding the 2.00 criterion indicating the nested nature of the data for these items. Overall, these statistics confirmed that, in general, responses within practice sites were more similar than would be expected by chance; therefore the clustered nature of the data should be taken into account when examining their factor structure.

**Multilevel Confirmatory Factor Analysis (MCFA):** All factor loadings estimated with the multilevel models were greater than the 0.40 criterion (Table 2). The between-practice site factor loadings ranged from 0.59 to 0.99 and the within-practice site factor loadings ranged from 0.45 to 0.99. The chi-square test (Table 3) was significant ( $p < 0.01$ ) as expected, but CFI was 0.97, above the 0.95 criterion. In addition, the RMSEA was 0.03, below the 0.06 criterion, indicating good fit. The within-practice site SRMR was 0.05, below the 0.08 criterion which indicated good fit, however the between-practice site SRMR was slightly above the cutoff at 0.10.

**Reliability**—All composites had acceptable (0.70 or above) individual level (internal consistency) reliability estimates, ranging from 0.77 to 0.89 (Table 4). Practice site level reliability was examined across the composites and global rating items by practice site size categories (1 clinician to 20 or more clinicians, Table 5). The practice site reliability estimates were acceptable for all sites with at least four clinicians. For sites with one clinician, only Access to Care had reliability above 0.70. The remaining reliabilities for practice sites with one clinician ranged from 0.40 (Courteous/Helpful Staff) to 0.69 (Overall Rating Item). For sites with 2–3 clinicians, both Access to Care and Courteous and Helpful Staff had reliability estimates above 0.70. The remaining reliabilities ranged from 0.58 (Recommend Doctor item) to 0.66 (Overall Rating item). The average number of respondents in 1-clinician and 2–3 clinician offices was less than 100, indicating that for these smaller sites it is necessary to have more respondents per practice site to increase reliability to acceptable levels.

**Spearman Correlations Among the Composites and Global Ratings**—All Spearman rank-order composite correlations were statistically significant ( $p < 0.01$ ), and none of the correlations exceeded the 0.80 criterion signaling potential multicollinearity (Table 4). The average individual level correlation among the composites was 0.30 (range:  $r = 0.25$  to  $r = 0.35$ ). The average practice site level correlation among the composites was

0.48 (range:  $r = 0.41$  to  $r = 0.57$ ). The lowest correlations at the individual and practice site levels were between Doctor Communication and Courteous/Helpful Staff (0.25 for individual and 0.41 for practice site level, respectively). The highest correlation at the individual level was between Access to Care and Doctor Communication ( $r = 0.35$ ). The highest correlation at the practice site level was between Access to Care and Courteous/Helpful Staff ( $r = 0.57$ ).

The Spearman correlations between the composites and the two global rating items were all statistically significant ( $p < 0.01$ ). For the Overall Doctor Rating item, the average individual level correlation with the composites was 0.38 (range:  $r = 0.22$  to  $r = 0.52$ ) and the average practice site level correlation was 0.50 (range:  $r = 0.34$  to  $r = 0.75$ ). For the Recommend Doctor item, the average individual level correlation with the composites was 0.38 (range:  $r = 0.29$  to  $r = 0.52$ ) while the average practice site level correlation was 0.57 (range:  $r = 0.43$  to  $r = 0.76$ ). The highest correlation with the global rating items was with the Doctor Communication composite and the Recommend Doctor item (0.52 at the individual level and 0.76 at the practice site level). Lastly, the Spearman correlations between the two global ratings were 0.47 and 0.76 at the individual and practice site levels, respectively.

## DISCUSSION

The CG-CAHPS Adult Visit survey is a publicly available, standardized tool to measure patients' experiences with outpatient medical offices. Demonstrating the psychometric properties of the survey is an important step for furthering its use. Overall, both the individual level and multilevel confirmatory factor analysis results provided support for the survey's three composites (Access to Care, Doctor Communication, Courteous/Helpful Staff) and two global rating items (Overall Doctor Rating, Recommend Doctor).

This study of a large number of practice sites and a large sample of patients provides support that the CG-CAHPS composites have acceptable individual-level internal consistency reliability as well as practice site level reliability. Practice level reliability is important because the survey is intended to provide information at the practice level, for public reporting of patient experience data, and to enable confidence in comparisons of data across sites. In our data set we found acceptable practice site level reliability for sites with at least four clinicians. The reliability stays relatively the same, and above 0.70, across sites with four to twenty or more clinicians (Table 5). Given that site-level reliability is a function of sample size, and the average sample size for practice sites with fewer than four clinicians was far less than those with four or more, these practice sites could achieve adequate site-level reliability by requiring responses from more respondents than were available in our data set.

The CG-CAHPS survey, in providing the patient's perspective, is critical for achieving the Institute of Medicine's aim of patient-centered care and for improving quality of care in outpatient medical offices. Numerous studies have linked patient experience data in various settings to better clinical outcomes, patient adherence to medications, patient retention in physicians' practices, and lower medical malpractice risk.<sup>17</sup> It is therefore important to have reliable and valid measures for assessing patient experience.

The associations between the composites and global rating items provide support for the construct validity of the CG-CAHPS measures. Doctor Communication had the strongest relationship with the global ratings, which is consistent with earlier studies that have shown doctor communication to be a key driver of patients' overall ratings of their doctor and their willingness to recommend their doctor.<sup>1, 2, 4</sup> The Courteous/Helpful Staff composite had the

weakest relationships with the global ratings suggesting that staff play less of a role in patients' global assessments of their doctors.

It should be noted that while there were a large number of practice sites included in our data set, they are not statistically representative of all medical offices in the U.S. because the data came from sites and states that voluntarily submitted their data to the CAHPS Database. Nevertheless, the analyses presented here represent one of the largest samples of medical offices studied and provide compelling support for the reliability, factor structure, and construct validity of the CG-CAHPS Adult Visit survey. Future research is needed to assess the associations of CG-CAHPS survey responses with clinical process measures and health outcomes.

## Acknowledgments

Work on the project described in this article was supported under contract (#HHS A290200710024C) and by cooperative agreements (#U18HS016978 and U18 HS016980) with the Agency for Healthcare Research and Quality. We thank Dale Shaller for facilitating access to data in the CAHPS Database.

## References

1. Solomon LS, Hays RD, Zaslavsky AM, et al. Psychometric properties of a group-level Consumer Assessment of Health Plans Study (CAHPS) instrument. *Med Care*. 2005; 43:53–60. [PubMed: 15626934]
2. Hargraves JL, Hays RD, Cleary PD. Psychometric Properties of the Consumer Assessment of Health Plans Study (CAHPS) 2.0 Adult Core Survey. *Health Serv Res*. 2003; 38(Part I):1509–1527. [PubMed: 14727785]
3. Hays RD, Brown J, Brown L, et al. Classical Test Theory and Item Response Theory Analyses of Multi-Item Scales Assessing Parents' Perceptions of Their Children's Dental Care. *Med Care*. 2006; 44(Suppl 3):S60–S68. [PubMed: 17060837]
4. Hays RD, Shaul JA, Williams V, et al. Psychometric Properties of the CAHPS™ 1.0 Survey Measures. *Med Care*. 1999; 37(Suppl 3):MS22–MS31. [PubMed: 10098556]
5. O'Malley AJ, Zaslavsky AM, Hays RD, et al. Exploratory Factor Analyses of the CAHPS Hospital Pilot Survey Responses across and within Medical, Surgical, and Obstetric Services. *Health Serv Res*. 2005; 40(Part II):2078–2095. [PubMed: 16316439]
6. Peterson RA. A Meta-Analysis of Variance Accounted for and Factor Loadings in Exploratory Factor Analysis. *Marketing Letters*. 2000; 11:261–275.
7. Brown, TA. *Confirmatory Factor Analysis for Applied Research*. New York: The Guilford Press; 2006.
8. Kenny, DA. [Accessed July 18, 2012.] Measuring Model Fit. 2011 Sep 4. Available at: <http://davidakenny.net/cm/fit.htm>
9. Hu LT, Bentler PM. Cutoff Criteria for Fit Indices in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. *Structural Equation Modeling*. 1999; 6:1–55.
10. Muthen, BO.; Muthen, LK. *Mplus: User's Guide*. Los Angeles, CA: Muthen & Muthen; 1998.
11. Julian M. The consequences of ignoring multilevel data structures in nonhierarchical covariance modeling. *Structural Equation Modeling*. 2001; 8:325–352.
12. Muthen BO, Satorra A. Complex Sample Data in Structural Equation Modeling. *Sociological Meth*. 1995; 25:267–316.
13. Nunnally, JC.; Bernstein, IH. *Psychometric Theory*. New York: McGraw Hill; 1994.
14. Raudenbush, SW.; Bryk, AS. *Hierarchical Linear Models*. 2. Thousand Oaks, Calif: Sage; 2002. p. 46equation 3.42
15. O'Brien RM. A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality and Quantity*. 2007; 41:671–690.



16. Nelson EC, Gentry MA, Mook KH, Spritzer KL, Higgins JH, Hays RD. How many patients are needed to provide reliable evaluations of individual clinicians? *Medical Care*. 2004; 42:259–266. [PubMed: 15076825]
17. Browne K, Roseman D, Shaller D, Edgman-Levitan S. Measuring Patient Experience As a Strategy for Improving Primary Care. *Health Affairs*. May; 2010 5(29):921–925. [PubMed: 20439881]

**Table 1**  
 Descriptive Statistics for CG-CAHPS Adult Visit Items with Associated Composite and Global Ratings

Composites, Items, and Global Ratings	Top Box	Mean	S.D.	Spearman Rank-Order Correlations		
				Associated Composite	Recommend Doctor	Overall Rating
<b>Access to Care Composite</b>						
Q6 Got appointment when needed care right away	64%	3.52	0.73	0.59	0.31	0.33
Q8 Got appointment for check-up or routine care	68%	3.60	0.66	0.59	0.30	0.31
Q10 Got answer to medical question same day	60%	3.46	0.76	0.50	0.30	0.33
Q13 Visit started < 15 minutes of appointment time	40%	3.14	0.87	0.40	0.23	0.26
<b>Doctor Communication Composite</b>						
Q18 Did Doctor explain things in a way easy to understand	92%	2.91	0.31	0.55	0.47	0.38
Q19 Did Doctor listen carefully	93%	2.93	0.29	0.57	0.53	0.39
Q21 Did Doctor give easy to understand instructions	91%	2.89	0.35	0.58	0.50	0.41
Q22 Did Doctor know important medical history	90%	2.89	0.36	0.49	0.45	0.41
Q23 Did Doctor show respect	95%	2.94	0.27	0.51	0.52	0.36
Q24 Did Doctor spend enough time	91%	2.89	0.35	0.53	0.48	0.40
<b>Courteous/Helpful Staff Composite</b>						
Q28 Clerks and receptionists helpful	88%	2.86	0.38	0.71	0.28	0.22
Q29 Clerks and receptionists treat with courtesy and respect	92%	2.91	0.31	0.71	0.26	0.18
<b>Single Item Global Ratings</b>						
Q25 Overall Doctor Rating	82%	9.27	1.33			
Q26 Recommend Doctor	90%	2.88	0.39			

Notes: Responses were from 450 Practice Sites and 21,318 respondents. The Spearman correlation with the associated composite is the item-to-composite corrected correlation.

Table 2

Confirmatory Factor Analysis (CFA) Results of the CG-CAHPS Adult Visit Survey

Composites and Items	Individual Level CFA		Practice Site Level Multilevel CFA		
	Factor Loading	ICC	Design Effect	Within Factor Loading	Between Factor Loading
<b>Access to Care Composite</b>					
Q6 Got appointment when needed care right away	0.82	0.08	4.66	0.81	0.98
Q8 Got appointment for check-up or routine care	0.81	0.07	4.06	0.81	0.98
Q10 Got answer to medical question same day	0.62	0.07	4.15	0.60	0.91
Q13 Visits started < 15 minutes of appointment time	0.46	0.11	6.24	0.45	0.59
<b>Doctor Communication Composite</b>					
Q18 Did Dr. explain things in a way easy to understand	0.75	0.01	1.51	0.75	0.99
Q19 Did Dr. listen carefully	0.85	0.02	1.70	0.85	0.98
Q21 Did Dr. give easy to understand instructions	0.77	0.01	1.60	0.77	0.96
Q22 Did Dr. know important medical history	0.65	0.02	2.02	0.65	0.94
Q23 Did Dr. show respect	0.81	0.02	1.70	0.81	0.99
Q24 Did Dr. spend enough time	0.73	0.02	1.93	0.73	0.91
<b>Courteous/Helpful Staff Composite</b>					
Q28 Clerks and receptionists helpful	0.91	0.05	3.27	0.74	0.99
Q29 Clerks and receptionists treat with courtesy and respect	0.81	0.04	2.95	0.99	0.82

**Table 3**

Fit Indices for Three-Factor Confirmatory Factor Analysis Models of the CG-CAHPS Adult Visit Survey

Fit Statistic	Individual Level CFA	Practice Site Level	Multilevel CFA
$\chi^2$	3129.6		2551.1
<i>df</i>	51		105
<i>CFI</i>	0.97		0.97
<i>RMSEA</i>	0.05		0.03
<i>SRMR</i>	0.04		--
<i>Within</i>	--		0.05
<i>Between</i>	--		0.10

Note: Chi-square statistics are significant,  $p < 0.01$ .

Table 4

Individual Level Reliability and Individual and Practice Site Level Spearman Rank-Order Correlations Among CG-CAHPS Adult Visit Composites and Global Ratings

Composites and Global Ratings	Composites and Global Rating Items				
	1	2	3	4	5
1. Access to Care Composite	(0.77)	0.45	0.57	0.42	0.52
2. Doctor Communication Composite	<b>0.35</b>	(0.89)	0.41	0.75	0.76
3. Courteous/Helpful Staff Composite	<b>0.29</b>	<b>0.25</b>	(0.85)	0.34	0.43
4. Overall Doctor Rating Item	<b>0.39</b>	<b>0.52</b>	<b>0.22</b>	-	0.76
5. Recommend Doctor Item	<b>0.34</b>	<b>0.52</b>	<b>0.29</b>	<b>0.47</b>	-

Notes: Individual-level Cronbach's alphas shown in parentheses ( ). The top right part of matrix displays Practice Site level correlations. Lower left, bolded part of matrix displays Individual level correlations. All Spearman correlations are statistically significant,  $p < 0.01$ .

**Table 5**  
Average Practice Site Level Reliabilities by Size for CG-CAHPS Adult Visit Survey

Composite/Item	Practice Site Size							
	1 Clinician	2–3 Clinicians	4–9 Clinicians	10–13 Clinicians	14–19 Clinicians	20 or more Clinicians		
<i>Number of Practice Sites</i>	36	84	211	53	49	18		
Access to Care Composite	0.77	0.70	0.83	0.91	0.89	0.81		
Doctor Communication Composite	0.62	0.62	0.78	0.79	0.81	0.79		
Courteous/Helpful Staff Composite	0.40	0.73	0.84	0.89	0.91	0.76		
Overall Doctor Rating Item	0.69	0.66	0.81	0.84	0.83	0.82		
Recommend Doctor Item	0.66	0.58	0.75	0.82	0.81	0.83		
Average respondent N	48	90	200	296	298	602		

*Note:* Practice site level reliabilities were calculated on the larger dataset consisting of 469 practice sites excluding 8 sites that were unable to be categorized by size.