Exploratory Factor Analyses of the CAHPS[®] Hospital Pilot Survey Responses across and within Medical, Surgical, and Obstetric Services

A. James O'Malley, Alan M. Zaslavsky, Ron D. Hays, Kimberly A. Hepner, San Keller, and Paul D. Cleary

Objectives. To estimate the associations among hospital-level scores from the Consumer Assessments of Healthcare Providers and Systems (CAHPS[®]) Hospital pilot survey within and across different services (surgery, obstetrics, medical), and to evaluate differences between hospital- and patient-level analyses.

Data Source. CAHPS Hospital pilot survey data provided by the Centers for Medicare and Medicaid Services.

Study Design. Responses to 33 questionnaire items were analyzed using patient- and hospital-level exploratory factor analytic (EFA) methods to identify both a patient-level and hospital-level composite structures for the CAHPS Hospital survey. The latter EFA was corrected for patient-level sampling variability using a hierarchical model. We compared results of these analyses with each other and to separate EFAs conducted at the service level. To quantify the similarity of assessments across services, we compared correlations of different composites within the same service with those of the same composite across different services.

Data Collection. Cross-sectional data were collected during the summer of 2003 via mail and telephone from 19,720 patients discharged from November 2002 through January 2003 from 132 hospitals in three states.

Principal Findings. Six factors provided the best description of inter-item covariation at the patient level. Analyses that assessed variability across both services and hospitals suggested that three dimensions provide a parsimonious summary of inter-item covariation at the hospital level. Hospital-level factor structures also differed across services; as much variation in quality reports was explained by service as by composite.

Conclusions. Variability of CAHPS scores across hospitals can be reported parsimoniously using a limited number of composites. There is at least as much distinct information in composite scores from different services as in different composite scores within each service. Because items cluster slightly differently in the different services, service-specific composites may be more informative when comparing patients in a given service across hospitals. When studying individual-level variability, a more differentiated structure is probably more appropriate.

Key Words. CAHPS Hospital survey, factor analysis, hospital quality, hierarchical model

The Consumer Assessments of Healthcare Providers and Systems (CAHPS[®]) Hospital survey was designed to enable patients, physicians, and payers to compare quality among hospitals and to facilitate quality improvement in hospitals. Efficient reporting of information about the quality of care and service at hospitals requires identifying the important dimensions of hospital care and reliably evaluating hospitals' performance in each dimension. To do so, surveys that assess the experiences of patients recently discharged from acute care hospitals are analyzed to characterize the dimensions which best summarize variation in patient responses. Such analyses are often performed using factor analyses; we argue that it is important that these analyses also be conducted at the hospital level and show that different results can be obtained from hospital- and patient-level analyses.

Initially, items for the CAHPS Hospital survey instrument were designed to address the Institute of Medicine's (IOM) domains of patient-centered care: respect for patient's values; preferences and expressed needs; coordination and integration of care; information, communication, and education; physical comfort; emotional support; involvement of family and friends; transition and continuity; and access to care (Goldstein et al. 2005). The cognitive testing phase of survey development eliminated many items and some domains (Levine, Fowler, and Brown 2005) for concepts that were too complicated, abstract, or subjective to support the development of unambiguous, easily understood items. This necessitated an exploratory approach to the factor analysis in order to identify composite items. The pilot survey included 33 questions to report on hospital care quality and four questions that elicited global ratings of care. The survey also included two open-ended questions concerning the hospital stay, 16 screener items that allowed respondents to skip the subsequent report items if they were ineligible to answer

Address correspondence to Paul D. Cleary, Ph.D., Professor, Department of Health Care Policy, Harvard Medical School, 180 Longwood Avenue, Boston, MA 02115-5899. A. James O'Malley, Ph.D., Assistant Professor of Statistics, and Alan M. Zaslavsky, Ph.D., Professor of Statistics, are with the Department of Health Care Policy, Harvard Medical School, Boston, MA. Ron D. Hays, Ph.D., and Kimberly A. Hepner, Ph.D., are with the RAND Corporation, Santa Monica, CA. San Keller, Ph.D., Principal Scientist, is with the American Institutes for Research (AIR), Chapel Hill, NC.

them, and 11 items on patient characteristics (to support case-mix adjustment models).

The 33 report questions were designed to collect information that is important to patients and discriminates among hospitals (Goldstein et al. 2005). Reports of CAHPS data typically present summaries or composites that average responses within groups of items determined by content relationships and/or empirical associations (AHCPR 1999; Hays et al. 1999; Zaslavsky et al. 2000; Bender and Garfinkel 2001; Marshall et al. 2001; Hargraves, Hays, and Cleary 2003). If several items are strongly associated and substantively similar, they can be combined to reduce the number of scores one must examine to understand variations in quality. Summary measures facilitate interpretation and use of data by consumers, clinicians, and others interested in monitoring and making decisions about health care (Hibbard et al. 2002; Hibbard, Stockard, and Tusler 2003). Factor analysis can be used to identify groups of empirically related items that are the product of the same latent variable.

One can assess patterns of associations at the individual level (identifying items that are scored similarly by patients) or hospital level (identifying items on which hospitals have similar scores). Each type of association can be informative for different purposes (Zaslavsky et al. 2000). To understand individual variations within hospitals (e.g., do men and women report different health care experiences?) patient-level associations are of greatest interest. On the other hand, to assess the relative performance of different hospitals, hospital-level associations are more relevant.

Comparisons of individual- and hospital-level analyses can also help address methodological issues in surveys. Correlations at the individual level might reflect individual patients' response tendencies (e.g., acquiescence bias), the common effects of patient characteristics on several kinds of experiences (e.g., cultural background), or response patterns related to the way the questions are presented or organized (e.g., context effects). A hospital-level analysis that removes the component of correlation because of patterns of patient responses that have nothing to do with quality of care might better reflect associations among aspects of hospital quality. Furthermore, the patient-level correlation analysis is confounded by different nonresponse patterns for the different items (because of skip instructions), while hospital-level mean scores can be calculated and correlated, even for pairs of items that are answered by nonoverlapping sets of patients.

We estimated hospital-level associations by fitting a two-level multivariate model to the hospital mean scores, in which these scores were modeled as

estimates, subject to error because of variation of individual patients, of the long-term population means for the corresponding hospitals.

METHODS

Data

Following the removal of 333 respondents with undetermined service or hospital affiliation, 19,720 patients in 132 hospitals were available for this analysis (Levine, Fowler, and Brown 2005). Two atypical hospitals with very few respondents were excluded, leaving 130 hospitals and 19,683 survey responses. We classified patients by the service (medical, surgical, or obstetric) in which they were treated using their diagnostic-related group (DRG) codes. Of the 130 hospitals that provided both surgery and medical services (7,904 and 7,183 survey responses, respectively), only 102 provided obstetrics care (4,596 responses). Thus, there were a total of 362 hospital-service units. The number of surgical, obstetric, and medical respondents per hospital had skewed distributions, with means (SDs) of 61 (52), 45 (42), and 55 (45), respectively.

Variables

We focus on the 33 report items because these are the measures that would be combined into multi-item composites. We did not adjust for patient covariates because preliminary analyses suggested that case-mix adjustment would have a minimal effect on factor analyses at the hospital level.

Factor Analyses

We conduct several factor analyses with different units of analysis: (1) a “consensus” hospital-level analysis in which each hospital service was treated as a distinct unit, revealing general patterns applying across all services; (2) a patient-level analysis; and (3) hospital-level analyses conducted separately for each service. We compared analyses (1) and (2) to determine whether patient- and hospital-level associations had similar structures, and compared factor structures across services in (3). Using these results, we grouped related items to define composites that parsimoniously summarized the full set of variables.

All individual- and hospital-level factor analyses were conducted on correlation matrices using the principal factor method with squared multiple correlations as initial communality estimates, Promax (oblique) rotation, and Kaiser normalization. Although hospital-level analysis also estimated the variances at the patient and hospital levels, this paper focuses on the correlation

structure; comparisons of the amount of variance explained at each level are discussed elsewhere (Keller et al. 2005).

For patient-level analyses, we used the maximum-likelihood correlation matrix calculated by *SAS PROC MI* under the missing-at-random (MAR) assumption, allowing us to use all available data despite the extensive structured item nonresponse (Little and Rubin 1987). In the case of structured missing data (i.e., because of item skip instructions), the missing values are regarded as not meaningful (e.g., the experience of getting help with bathing, washing, or keeping clean of a patient who reported having no need for such help) and so the MAR assumption is not relevant to the analysis. Therefore, the imputation procedure relies on no assumptions that are inconsistent with the patterns observed in the data and so may be used to obtain a maximum likelihood estimate of a covariance matrix that is consistent with what is observed. This allows us to use exploratory factor analysis with all available data to describe the relationships at the patient level.

For the hospital- and hospital-by-service-level analyses, we used a two-level hierarchical model to estimate the unit-level correlation structure, removing the component of variance because of error at the patient-level (i.e. sampling variability within hospitals) prior to evaluating the between-hospital correlation matrix (see a brief technical description of this analytic approach in the Appendix). We adjusted for service main effects in the consensus hospital-level analysis.

In the hospital-level factor analysis, an item is initially assigned to a factor on which it has a factor pattern loading (standardized regression coefficient) greater than 0.40 or, if it has multiple high loadings, to the factor with the highest loading. Because the rotated factor pattern matrix maximizes the separation between high and low loadings it can clarify which items load uniquely on which factors. Both substantive considerations and empirical criteria (e.g., the variation explained by each factor) were used to select the number of factors and to make final group assignments. In the patient-level factor analysis, the loading criterion was set to 0.30 because sampling variation led to lower correlations for some items (Child 1970).

Correlation Analysis of Composites within and across Services

Using composite scores (unweighted means of hospital mean scores on items) based on groupings determined from the factor analyses, we calculated correlations of the composite scores both within and across services to determine the extent to which different services within the same hospital were evaluated

similarly by patients. These correlations were corrected for sampling variability through the hierarchical model. To determine the extent to which hospitals that perform well in one quality dimension perform well in other dimensions and the extent to which performance on different services within hospitals was correlated, we calculated the average correlation between different composites for the same service, between the same composite in different services, and between different composites in different services.

RESULTS

Hospital-Level Consensus Factor Analysis

Four hospital-level, unrotated consensus factors had eigenvalues (20.48, 4.67, 1.35, 1.03) greater than the average eigenvalue of 0.95 (the mean eigenvalue does not equal 1 because the unique variation of each item is subtracted prior to analysis), and thus satisfied Guttman's criteria (Guttman 1954) for factor selection. We report a three-factor solution because only a single variable had a large loading on the fourth factor.

We labeled the three factors: (1) physician treatment, discharge information, and pain management (for short, "doctor"), (2) nurse treatment and support, medication information (for short, "nursing"), and (3) environment (Table 1). Two items (tests without pain, bathroom help) loaded highly on both the doctor and nursing factors, while several other items (e.g., the pain and medication-related items) loaded highly on one of these factors and moderately on the other. The environment factor groups several items related to the room and the initial meetings of staff with the patient. The correlations between the consensus factors ranged between 0.46 and 0.50 (Table 1) implying moderate overlap among the factors. Several items had low communalities (<0.70), suggesting that they were relatively weakly related to any of the other items. Living Will had by far the lowest communality and does not load above our threshold for grouping with any factor. Other items with low communalities were: discharge medicine information, introductions, room quiet, tests without pain, and privacy.

Patient-Level Factor Analysis

The number of patient-level factors was determined by considering the eigenvalues and the interpretability of the factor pattern matrix of the rotated factors. Because the largest eigenvalues were 10.81, 1.51, 1.40, 0.89, 0.68, 0.48, 0.41 and the average eigenvalue was 0.45, we selected six factors based on Guttman's criteria. When we used a 0.30 threshold for loadings, all but four

Table 1: Hospital-Level Consensus Factor Analysis

	<i>Factor</i>			
	<i>Doctor</i>	<i>Nursing</i>	<i>Environment</i>	
<i>Correlation with other factors</i>				
Doctor	1.00	0.46	0.46	
Nursing	0.46	1.00	0.50	
Environment	0.46	0.50	1.00	
Variance explained eliminating other factors	8.60	6.05	1.68	
<i>Item</i>	<i>Standardized Regression Coefficients</i>			<i>Communality</i>
MD respect (Q11)	0.96	− 0.06	0.05	0.92
MD listen (Q12)	0.98	− 0.10	0.02	0.89
MD explanations (Q13)	0.99	− 0.11	0.02	0.91
MD enough time (Q14)	0.95	0.04	− 0.02	0.92
Involved in decisions (Q25)	0.82	0.23	0.03	0.93
Response to pain (Q31)	0.65	0.38	0.06	0.85
Pain controlled (Q32)	0.57	0.35	0.10	0.73
Pain help (Q33)	0.67	0.33	0.13	0.90
Tests without pain (Q35)	0.53	0.49	− 0.13	0.66
Purpose of medicines (Q38)	0.68	0.35	− 0.07	0.75
Delays in admission (Q43)	1.03	− 0.28	0.04	0.91
Discharge activity information (nQ47)	0.98	− 0.14	0.02	0.88
Help after discharge (nQ48)	0.71	0.20	0.11	0.78
Discharge symptoms information (nQ49)	0.97	− 0.12	0.01	0.85
Discharge medicine information (nQ51)	0.53	0.33	0.06	0.60
Nurse respect (Q4)	− 0.05	0.89	0.14	0.90
Nurse listen (Q5)	0.07	0.77	0.23	0.89
Nurse explanations (Q6)	0.41	0.67	0.02	0.90
Nurse enough time (Q7)	0.31	0.60	0.24	0.90
Nurse help (Q9)	0.10	0.81	0.15	0.91
Help bathing (Q20)	0.18	0.79	− 0.03	0.75
Bathroom help (Q22)	0.60	0.46	0.02	0.84
Privacy (Q24)	0.37	0.76	− 0.45	0.68
Help visitors (Q27)	0.31	0.51	0.34	0.88
Names of medicines (Q37)	0.31	0.74	− 0.09	0.77
Questions about other medicines (Q39)	− 0.55	1.01	0.02	0.82
Questions about medicine allergies (Q40)	− 0.43	0.95	0.04	0.74
Explain medicine side effects (Q41)	0.21	0.64	0.23	0.82
Room temperature (Q16)	0.38	− 0.16	0.75	0.81
Room clean (Q17)	0.00	0.46	0.53	0.74
Room quiet (Q18)	− 0.12	0.10	0.80	0.65
Introductions (Q28)	0.16	0.31	0.49	0.63
Living Will (Q44)	0.29	0.07	0.38	0.38

Boldface entries are factor loadings for the hospital-level consensus composites items. Because only the unique amount of explained variation is reported, the total explained variation of the rotated factors can be less than the sum of the eigenvalues. Because the rotated factors need not be orthogonal, the loadings can exceed 1.0 in magnitude.

Table 2: Patient-Level Factor Analysis

	Factor					Consensus Factor
	MDCom	PainCont	Discharge	NursCom	PhysCom	ComMeds
<i>Correlation with other factors</i>						
MDCCom	1.00	0.48	0.39	0.55	0.53	0.49
PainCont	0.48	1.00	0.29	0.62	0.63	0.48
Discharge	0.39	0.29	1.00	0.36	0.53	0.46
NursCom	0.55	0.62	0.36	1.00	0.74	0.55
PhysCom	0.53	0.63	0.53	0.74	1.00	0.58
ComMeds	0.49	0.48	0.46	0.55	0.58	1.00
Variance explained eliminating other factors	1.71	0.89	1.07	0.84	0.97	1.46

Item	Standardized Regression Coefficients					MD services
MD respect (Q11)	0.83	0.03	- 0.03	0.01	- 0.01	- 0.05
MD listen (Q12)	0.88	0.01	- 0.02	0.01	- 0.02	- 0.01
MD explanations (Q13)	0.80	- 0.02	0.01	- 0.01	0.01	0.04
MD enough time (Q14)	0.72	- 0.01	0.02	0.01	0.09	0.00
Involved in decisions (Q25)	0.30	0.06	0.09	0.07	0.18	0.16
Response to pain (Q31)	- 0.02	0.65	0.01	0.11	0.12	0.02
Pain controlled (Q32)	0.02	0.75	0.01	- 0.05	0.07	0.00
Pain help (Q33)	0.04	0.74	0.03	0.05	0.06	0.02
Tests without pain (Q35)	0.09	0.20	0.00	0.00	0.18	0.05
Purpose of medicines (Q38)	0.06	- 0.06	0.01	0.04	0.03	0.75
Delays in admission (Q43)	0.10	0.07	0.05	- 0.04	0.17	- 0.02
Discharge activity information (nQ47)	0.01	- 0.01	0.69	- 0.03	- 0.01	- 0.06
Help after discharge (nQ48)	0.00	0.03	0.47	0.01	0.04	0.04

continued

Item	Standardized Regression Coefficients					
Discharge symptoms information (nQ49)	0.00	-0.03	0.70	0.00	-0.02	-0.01
Discharge medicine information (nQ51)	-0.02	0.04	0.42	-0.01	0.00	0.08
Nurse respect (Q4)	-0.03	0.05	-0.02	0.75	0.06	-0.01
Nurse listen (Q5)	0.01	0.03	-0.01	0.78	0.05	0.01
Nurse explanations (Q6)	0.10	-0.06	0.03	0.66	0.06	0.07
Nurse enough time (Q7)	0.04	0.00	0.01	0.56	0.24	-0.01
Nurse help (Q9)	-0.04	0.12	-0.02	0.42	0.31	0.00
Help bathing (Q20)	-0.02	-0.06	0.04	0.06	0.74	-0.02
Bathroom help (Q22)	-0.05	0.00	0.03	0.11	0.69	-0.01
Privacy (Q24)	0.10	0.09	0.01	0.00	0.44	0.08
Help visitors (Q27)	0.12	0.12	0.01	0.15	0.35	0.02
Names of medicines (Q37)	0.04	-0.04	0.02	0.02	0.02	0.73
Questions about other medicines (Q39)	-0.05	0.04	-0.03	-0.01	-0.01	0.76
Questions about medicine allergies (Q40)	-0.06	0.05	-0.01	-0.02	0.00	0.72
Explain medicine side effects (Q41)	0.03	0.01	0.10	0.01	0.08	0.58
Room temperature (Q16)	0.04	0.07	-0.07	-0.07	0.48	0.00
Room clean (Q17)	-0.02	-0.03	-0.04	0.00	0.61	0.05
Room quiet (Q18)	0.03	0.01	-0.06	-0.05	0.54	0.01
Introductions (Q28)	0.10	0.08	0.01	0.12	0.16	0.23
Living Will (Q44)	-0.01	0.06	0.19	0.02	-0.09	0.03

Boldface entries are factor loadings for the patient-level composites items. Items are ordered by the factors in the consensus hospital-level factor analysis.

items grouped with factors; one of the remaining items (tests without pain) was subsequently associated with the pain control factor, leaving three items out of the factor structure. To facilitate comparison, the rows of the resulting patient-level factor analysis (Table 2) are ordered according to the consensus hospital-level factor analysis results. The correlations between the factors of the patient-level analysis ranged from 0.29 to 0.74 with most above 0.45. Compared with the hospital-level analysis, the patient-level factor analysis more often grouped adjacent items in the same factor: items 16–18 in the physical comfort factor, items 11–14 in the MD communication factor, items 37–41 in the medicine communication factor, items 4–7 in the nurse communication factor, items 31–33 in the pain control factor, and items 47–49 in the discharge information factor. Within several of these factors, the largest loadings were for groups of consecutive items. Thus, the correlations at the patient level could be due in part to either the likeness of the topic or the adjacency of items concerning the same topic in the survey.

In some cases, the hospital-level consensus factor structure groups items in a way that collapses together several of the patient-level groups. Items that load on the hospital-level doctor factor combine the patient-level MD communication, pain control, and discharge information groups. Similarly, the hospital-level nursing factor combines the nursing-related component of the patient-level physical comfort factor, the communication about medication factor, and the nurse communication factor. The items loading on the environment factor essentially constitute the remainder of the physical comfort factor.

Hospital-Level Analyses for Each Service

Five factors for the surgery and obstetric services and four factors for medical had eigenvalues exceeding the mean. However, based on arguments similar to those considered in the consensus analysis (e.g., the presence of single-item factors) and because these subset analyses were based on fewer observations and therefore are more susceptible to model misspecification, we report fewer (three) rather than more factors. These accounted for nearly 80 percent of the variance explained for each of the three services (Table 3).

The first factor for surgical services approximately combines the nursing and environment factors of the consensus analysis. The second factor resembles the doctor factor, while the third factor is a grouping of items related to discharge information and nonclinical items (Privacy, Helps Visitors, Living Will).

Table 3: Hospital-Level Service Specific Analyses

Correlation with other factors	Surgery Factors			Obstetrics Factors			Medical Factors		
	MD								
	Services	Discharge	RN+Env	RN+MD	Environ	Privacy	DirCare	MedPainEnv	Discharge
Factor 1	1.00	0.57	0.62	1.00	0.60	0.52	1.00	0.65	0.48
Factor 2	0.57	1.00	0.61	0.60	1.00	0.37	0.65	1.00	0.59
Factor 3	0.62	0.61	1.00	0.52	0.37	1.00	0.48	0.59	1.00
Variance explained eliminating other factors	3.33	3.03	4.55	7.41	3.05	2.10	3.97	3.97	2.98

Item	Standardized Regression Coefficients			Standardized Regression Coefficients			Standardized Regression Coefficients			Consensus Factor
MD respect (Q11)	0.98	-0.22	0.13	0.73	0.01	0.21	0.94	0.12	-0.10	MD Services
MD listen (Q12)	0.92	-0.11	0.11	0.73	0.14	0.07	0.96	0.08	-0.22	
MD explanations (Q13)	0.91	-0.09	0.12	0.74	0.06	0.10	0.98	0.09	-0.21	
MD enough time (Q14)	0.90	0.07	-0.01	1.02	0.01	-0.15	0.94	-0.02	0.05	
Involved in decisions (Q25)	0.49	0.16	0.43	0.85	0.04	0.14	0.38	0.36	0.30	
Response to pain (Q31)	0.41	0.07	0.56	0.61	0.22	0.22	0.24	0.73	-0.04	
Pain controlled (Q32)	0.49	-0.01	0.47	0.55	0.26	0.22	0.51	0.59	-0.23	
Pain help (Q33)	0.37	0.30	0.42	0.64	0.27	0.19	0.31	0.54	0.18	
Tests without pain (Q35)	0.46	0.60	-0.16	1.16	-0.50	-0.19	0.33	0.51	0.13	
Purpose of medicines (Q38)	0.04	0.30	0.49	0.55	0.09	0.23	-0.01	0.81	0.14	
Delays in admission (Q43)	0.80	0.24	-0.41	0.69	0.14	-0.01	0.80	-0.33	0.28	
Discharge activity information (nQ47)	0.01	1.00	-0.23	0.76	0.06	-0.01	0.07	0.26	0.63	
Help after discharge (nQ48)	0.13	0.73	0.13	0.81	0.18	-0.19	0.26	-0.06	0.78	
Discharge symptoms information (nQ49)	0.05	0.94	-0.13	0.87	-0.05	-0.04	-0.16	-0.16	1.01	

Discharge medicine information (nQ51)	0.16	0.62	0.09	0.05	-0.18	0.80	-0.21	0.16	0.73	
Nurse respect (Q4)	0.25	0.25	0.57	0.67	0.27	0.17	0.37	0.35	0.38	
Nurse listen (Q5)	0.31	0.17	0.60	0.51	0.42	0.21	0.40	0.36	0.34	
Nurse explanations (Q6)	0.17	0.55	0.38	0.74	0.24	0.09	0.52	0.18	0.43	
Nurse enough time (Q7)	0.20	0.41	0.48	0.76	0.22	0.08	0.32	0.28	0.54	
Nurse help (Q9)	0.10	0.27	0.70	0.60	0.32	0.14	0.26	0.53	0.30	RN Services
Help bathing (Q20)	0.28	0.15	0.54	0.79	0.24	-0.24	0.55	0.02	0.45	
Bathroom help (Q22)	0.16	0.40	0.50	0.81	0.17	0.02	0.34	0.22	0.41	
Privacy (Q24)	0.27	0.43	0.15	-0.04	-0.16	0.89	0.56	0.23	0.01	
Help visitors (Q27)	0.05	0.52	0.47	0.59	0.45	0.03	0.23	0.64	0.18	
Names of medicines (Q37)	-0.05	0.40	0.54	0.85	-0.08	0.09	-0.02	0.80	0.18	
Questions about other medicines (Q39)	-0.01	0.03	0.86	0.28	0.55	-0.07	-0.12	0.88	0.10	
Questions about medicine allergies (Q40)	-0.11	0.00	0.87	0.31	0.56	-0.09	-0.14	0.95	0.05	
Explain medicine side effects (Q41)	0.02	0.11	0.83	0.53	0.51	-0.08	0.13	0.77	0.06	
Room temperature (Q16)	-0.06	-0.01	0.64	-0.10	0.76	0.23	0.12	0.72	-0.06	Environment
Room clean (Q17)	-0.07	-0.08	0.96	0.22	0.77	-0.20	0.37	0.56	0.07	
Room quiet (Q18)	-0.07	-0.26	0.90	-0.26	1.08	-0.24	0.33	0.59	-0.21	
Introductions (Q28)	0.32	0.03	0.52	0.08	0.46	0.46	-0.10	0.70	0.21	
Living Will (Q44)	-0.47	0.67	0.30	-0.07	-0.01	0.75	-0.17	0.29	0.62	

Boldface entries are factor loadings for the hospital-level service specific composites items. Items are ordered by the factors in the consensus hospital-level factor analysis.

The factor structure for obstetric care is dominated by the first factor, which combines the items grouped by the doctor and nursing consensus factors. The second factor groups the environment and medication items, and the third groups the items on Discharge Medicine Information, Privacy, and Living Will.

For medical services, most items on doctor and nursing services load on a single factor. Items on pain control, medications, and the environment load on a second factor. The third factor groups discharge-related items, as for surgical services.

The correlations between the service-specific factors are between 0.48 and 0.62 in most instances. Thus, as in the consensus analysis, much of the observed variation in all of the groups of items appears to be explained by one general quality dimension.

Correlations of Composites between Different Services

We compared the correlation between the hospital-level consensus composites within and between services. Correlations between composites on the same service were summarized by the average correlation over all pairs of composites for that service. Correlations between services of the same composites and of different composites were summarized by the average correlation over the respective sets of composites for each pair of services.

The mean within-service correlations of the consensus factors for surgery, obstetrics, and medical were 0.76, 0.73, and 0.85, respectively. These compared with mean correlations of 0.55, 0.71, and 0.51 for the same composite across different services; and 0.48, 0.59, and 0.45 across different composites in different services, for the surgery–obstetrics, surgery–medical, and obstetrics–medical pairs of services, respectively.

These results suggest that surgery and medical are the most similar services and that medical and obstetrics are the most distinct services. Correlations between composites within the same service are higher than the correlations between the same composites in different services. This suggests that differences between services are more pronounced than differences between dimensions of quality, or in other words that the services themselves are the most important dimensions of quality. As expected, correlations between different composites across different services are smaller than within service or within composites.

DISCUSSION

The desirability and method of combining responses to a patient survey depends on the purposes for which the data are to be used. For quality improvement efforts, it is often desirable to know the response distribution to specific questions. The majority of the questions in the CAHPS Hospital survey were selected because they ask about aspects of care that are important to patients, and these specific aspects might be more “actionable” than general assessments such as the global ratings of care received from doctors or nurses questions.

To develop more parsimonious measures that simplify the reporting and interpretation of variability, however, it may be useful to combine items (Veroff et al. 1998; McGee et al. 1999; Sofaer 1999). Groupings identified by the individual-level factor analyses suggest ways in which the items can be combined to compare reports across types of patients (e.g., men and women). Hospital-level analyses may yield a different way of combining responses that is more appropriate to describing interhospital variability in quality.

When we examined variability across both services and hospitals together, we found that three dimensions explained a substantial amount of the interunit variability. Although this was also true within each service (surgical, obstetric, and medical), we found somewhat different patterns in each service, suggesting that the relationships of the various aspects of care are different. The questions with the most distinct patterns across services were the pain control, medication, and discharge information items. This is reasonable because pain management and discharge planning are quite different for medical, surgical, and obstetric patients. For example, discharge information after surgery or medical treatment is focused on continuity with postdischarge outpatient treatment and recovery, typically involving management of some therapeutic or analgesic medications. The concerns of obstetric patients, however, are usually focused on care for the newborn, a topic that is not specifically addressed in the survey. Similarly, the most salient issues in pain management for obstetric patients concern administration of analgesia in accord with patient preferences during childbirth, while surgical patients are not likely to be aware of operative anesthesia, but are more concerned with management of postoperative pain.

Nursing services and environment are closely related for surgery, but more sharply distinguished from doctor items: the surgeon’s role in performing surgery is quite distinct from that of nurses, who ensure that the patient is comfortable before and after surgery. Conversely in obstetrics the doctor and

nursing items group together, reflecting the greater sharing of roles between doctors and nurses during childbirth. Similarly, in medical services, doctors and nurses may have similar effects on patient experiences. Patients receiving general medical services may be in discomfort and thus quite sensitive to the attention they received from personnel at the hospital, and so it is reasonable that items related to pain relief, treatment, and comfort group together. The items in the environment consensus factor also were grouped for each of the three services, but only constituted a distinct dimension of quality for obstetrics. Although the number of hospitals included in this study was not sufficient to allow us to generalize with confidence about these differences among services, the differences we found suggest that the characteristics of hospital quality are structured differently for surgical, general medical, and obstetric services.

Correlations between different consensus composites within the same service were higher than correlations between the same composites in different services, indicating that differences between services are just as important to recognize as differences between items. The generally highly positive correlations across services and dimensions of quality suggest, however, that there is a general dimension to hospital quality: high-quality hospitals may be expected to perform well across all services and dimensions.

The patient-level factor analysis yields a more differentiated factor structure than that estimated at the hospital level. Furthermore, when we constrained the patient-level factor analysis to three or four factors (not shown), the dimensions of quality did not agree with those found at the hospital level. It is likely that different patients have different constellations of experiences because such experiences are determined in part by individual conditions, treatments, preferences, and other characteristics. It is also possible that methodological aspects of the survey such as the positioning of questions in the questionnaire and patterns of missing data because of skip patterns affected the pattern of individual level correlations. However, the patient-level pain control, communication about medications, and discharge information factors remained intact in the hospital-level factor solutions. The results of these factor analyses, in combination with the results of other analyses (Keller et al. 2005), were used to make recommendations of items to drop from the CAHPS Hospital survey.

There are several limitations to this study. The data analyzed came from only 130 hospitals and we analyzed the relationships among 33 different items (for each of the three services). Thus, the ratio of observations to variables analyzed was not as high as desirable. Furthermore, some hospitals had limited numbers of respondents, which together with structured nonresponse

across items made direct estimates of the sampling covariance matrices relatively unstable. With more hospitals and larger sample sizes within hospitals, the estimates of the between-hospital covariance matrices would be more robust and less reliant on modeling assumptions.

The analyses reported herein are based on data from only three states. Therefore, the results do not necessarily reflect relationships that would be found in the U.S. population as a whole. Furthermore, we have not accounted for the effect of state or the effect of patient-level covariates, although the latter are likely to have minimal effect on the hospital-level factor analyses.

To provide the most effective information to consumers, clinicians, and managers, it is important that reports provide the most appropriate and accurate information and be designed in a way that eases interpretation. The analyses presented herein suggest that the number and nature of composites may vary, depending on the intended use and audience. The six composites identified in the patient-level analyses may be the most appropriate method to represent variability across groups of individuals. For comparisons of a particular service (medical, surgical, or obstetric) across hospitals, service-specific composites may be most informative. To parsimoniously describe interhospital, between-service variability, the three composites may be most efficient. These analyses should be repeated with a larger and more representative set of hospitals before final decisions are made about composite construction.

ACKNOWLEDGMENTS

The CAHPS II project is funded by the Agency for Healthcare Research and Quality (AHRQ; grant number 5 U18 HS00924) and the Centers for Medicare and Medicaid Services through cooperative agreements with Harvard Medical School, RAND, and the American Institutes for Research. User support is provided through a contract with Westat. Additional information about the study can be obtained by calling the AHRQ Clearinghouse at 800-358-9295. The authors thank project officers Chris Crofton, Chuck Darby, Beth Kosiak, and MaryBeth Farquahr, and colleagues at the Centers for Medicare and Medicaid Services, Elizabeth Goldstein and Thomas Reilly for their active participation and helpful suggestions throughout the project and members of the CAHPS consortium for their role in the design and implementation of the data collection activities and helpful comments on an earlier draft of this manuscript.

REFERENCES

- AHCPR. 1999. *CAHPS 2.0 Survey and Reporting Kit*. Rockville, MD: Agency for Health Care Policy and Research.
- Ansari, A., K. Jedidi, and L. Dube. 2002. "Heterogeneous Factor Analysis Models: A Bayesian Approach." *Psychometrika* 67: 49–77.
- 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 Service Research* 36 (3): 489–508.
- Child, D. 1970. *The Essentials of Factor Analysis*. New York: Holt, Rinehart and Winston.
- Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. "Maximum Likelihood from Incomplete Data via the EM Algorithm." *Journal of the Royal Statistic Society, Series B* 39: 1–22.
- Duncan, T. E., S. C. Duncan, A. Alpert, H. Hops, M. Stoolmiller, and B. O. Muthen. 1997. "Latent Variable Modeling of Longitudinal and Multilevel Substance Use Data." *Multivariate Behavioral Research* 32: 275–318.
- Goldstein, L., C. Crofton, S. Garfinkel, and C. Darby. 2005. "Why Another Patient Survey Matters." *Health Services Research* 40. DOI: 10.1111/j.1475-6773.2005.00477.x. Available at www.blackwell-synergy.com.
- Guttman, L. 1954. "Some Necessary Conditions for Common-Factor Analyses." *Psychometrika* 19: 149–61.
- Hargraves, J. L., R. D. Hays, and P. D. Cleary. 2003. "Psychometric Properties of the Consumer Assessment of Health Plans (CAHPS[®]) 2.0 Adult Core Survey." *Health Service Research* 38 (6): 1509–27.
- Hays, R. D., J. A. Shaul, V. S. L. Williams, J. S. Lubalin, H.-K. L.D., S. F. Sweeny, and P. D. Cleary. 1999. "Psychometric Properties of the CAHPS 1.0 Survey Measures." *Medical Care* 37 (3, suppl): MS22–31.
- Hibbard, J. H., P. Slovic, E. Peters, and M. L. Finucane. 2002. "Strategies for Reporting Health Plan Performance Information to Consumers: Evidence from Controlled Studies." *Health Services Research* 37 (2): 291–313.
- Hibbard, J. H., J. Stockard, and M. Tusler. 2003. "Does Publicizing Hospital Performance Stimulate Improvement Efforts?" *Health Affairs* 22 (2): 84–94.
- Keller, S., A. J. O'Malley, R. D. Hays, R. Matthew, A. M. Zaslavsky, K. A. Hepner, and P. D. Cleary. 2005. "Methods Used to Streamline the CAHPS[®] Hospital Survey." *Health Services Research*. DOI: 10.1111/j.1475-6773.2005.00478.x. Available at www.blackwell-synergy.com.
- Levine, R., J. Fowler, and J. Brown. 2005. "Cognitive Interviews and Design of the CAHPS Hospital Survey." *Health Services Research*. DOI: 10.1111/j.1475-6773.2005.00472.x. Available at www.blackwell-synergy.com.
- Linn, R. L., J. A. Centra, and L. Tucker. 1975. "Between, within, and Total Group Factor-Analyses of Student Ratings of Instruction." *Multivariate Behavioral Research* 10 (3): 277–88.
- Little, R. J. A., and Donald B. Rubin. 1987. *Statistical Analysis with Missing Data*. New York: John Wiley & Sons.

- Marshall, G. N., L. S. Morales, M. Elliot, K. Spritzer, and R. D. P. A. Hays. 2001. "Confirmatory Factor Analysis of the Consumer Assessment of Health Plans Study (CAHPS) 1.0 Core Survey." *Psychological Assessment* 13 (2): 216–29.
- McGee, J., D. Kanouse, S. Sofaer, J. Hargraves, E. Hoy, and S. Kleinman. 1999. "Making Survey Results Easy to Report to Consumers: How Reporting Needs Guided Survey Design in CAHPS." *Medical Care* 37 (3): MS32–40.
- Muthen, B. O. 1991. "Multilevel Factor Analysis of Class and Student Achievement Components." *Journal of Education Measurement* 28: 338–54.
- . 1993. "Multilevel Covariance Structure Analysis." *Sociological Research Methods and Research* 22: 376–98.
- Nelder, J. A., and D. Pregibon. 1987. "An Extended Quasi-Likelihood Function." *Biometrika* 74: 221–32.
- O'Malley, A. J., and A. M. Zaslavsky. 2003. "Cluster Level Factor Analysis for Survey Data with Structured Nonresponse." *Proceedings of the Joint Statistical Meetings, American Statistical Association*.
- . 2005. "Variance–Covariance Functions for Domain Means of Ordinal Survey Items." *Survey Methodology*, 31 (2).
- Sofaer, S. 1999. "Strategic Considerations in Providing Quality Information to Medicare Beneficiaries in a Changing Decision Environment." The Commonwealth Fund.
- Veroff, D. R., P. M. Gallagher, V. Wilson, M. Uyeda, J. Merselis, E. Guadagnoli, S. Edgman-Levitan, A. Zaslavsky, S. Kleinmann, and P. D. Cleary. 1998. "Effective Reports for Health Care Quality Data: Lessons from a CAHPS Demonstration in Washington State." *International Journal of Quality in Health Care* 10 (6): 555–60.
- 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: 162–74.

SUPPLEMENTARY MATERIAL

The following supplementary material for this article is available online:

APPENDIX S1. Hospital level covariance modeling.