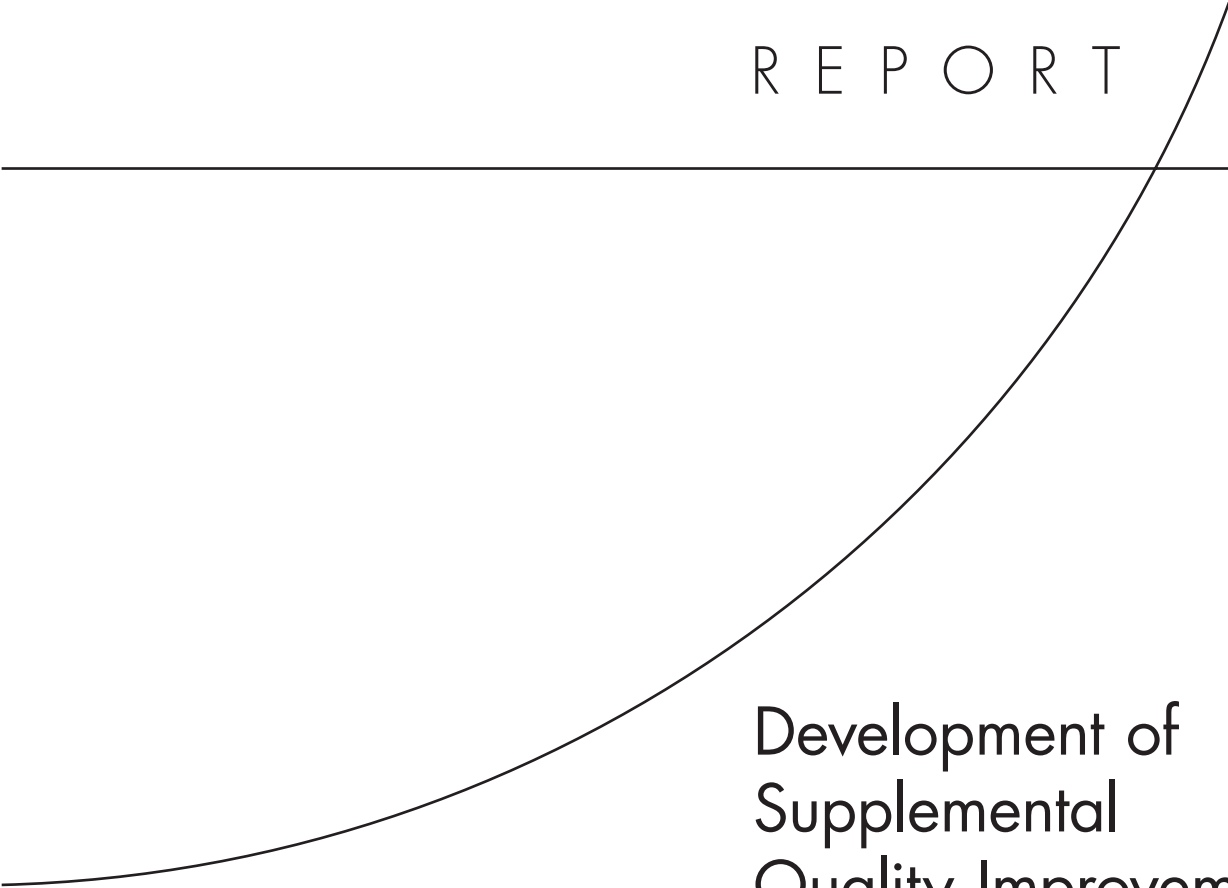


R E P O R T

---



# Development of Supplemental Quality Improvement Items for the Consumer Assessment of Healthcare Providers and Systems (CAHPS)

Denise D. Quigley, Donna O. Farley, Julie A. Brown,  
Marc N. Elliott, Han de Vries, Ron D. Hays

Sponsored by the Agency for Healthcare Research and Quality

The research described in this report was carried out in RAND Health, a division of the RAND Corporation. This work was sponsored by the Agency for Healthcare Research and Quality.

**Library of Congress Cataloging-in-Publication Data**

Development of supplemental quality improvement items for the Consumer Assessment of Healthcare Providers and Systems (CAHPS) / Denise D. Quigley ... [et al.].

p. ; cm. — (TR-336)

Includes bibliographical references.

ISBN-13: 978-0-8330-4002-2 (pbk. : alk. paper)

1. Insurance, Health—United States—Evaluation. 2. Managed care plans (Medical care)—United States—Evaluation. 3. Patient satisfaction—United States. I. Quigley, Denise D. II. Rand Corporation. III. Series: Technical report (Rand Corporation) ; TR-336.

[DNLM: 1. Consumer Assessment of Healthcare Providers and Systems (U.S.) 2. Health Care Surveys—United States—Technical Report. 3. Insurance, Health—United States—Technical Report. 4. Patient Satisfaction—United States—Technical Report. 5. Quality Assurance, Health Care—methods—United States—Technical Report. W 84.3 D489 2006].

RA412.2D48 2006

362.1'04258—dc22

2006024148

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

**RAND®** is a registered trademark.

A profile of RAND Health, abstracts of its publications, and ordering information can be found on the RAND Health home page at [www.rand.org/health](http://www.rand.org/health).

© Copyright 2006 RAND Corporation

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from RAND.

Published 2006 by the RAND Corporation

1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138

1200 South Hayes Street, Arlington, VA 22202-5050

4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213

RAND URL: <http://www.rand.org/>

To order RAND documents or to obtain additional information, contact

Distribution Services: Telephone: (310) 451-7002;

Fax: (310) 451-6915; Email: [order@rand.org](mailto:order@rand.org)

## **PREFACE**

The Consumer Assessment of Healthcare Providers and Systems (CAHPS®) study was initiated by the Agency for Healthcare Research and Quality (AHRQ) in 1995 to establish survey and reporting products that provide consumers information on the performance of health plans and providers, as judged by other consumers who have used the health plan. It was learned during the first CAHPS development work that various stakeholders have very different perspectives and uses for CAHPS. AHRQ initiated a second CAHPS project (CAHPS® II) in 2002, working with a consortium of RAND Corporation, American Institutes for Research, the Harvard Medical School, and Westat, an employee-owned contract research organization founded in 1961 headquartered in the Washington, DC, area. A goal of the CAHPS II is to address concerns expressed by health plans and health care providers that CAHPS does not provide them the information they need to improve performance in areas of importance to consumers.

Using a structured consensus process, called a Delphi method, whereby experts are asked to provide input independently to a set of questions, and then asked to respond individually to the summarized results, which are then fed back to the experts in an iterative process, RAND worked with health plans to identify the specific issues of concern regarding the usability of CAHPS results for quality improvement actions. Results of interviews with health plans confirmed their desire to have supplemental items available that they could use to collect more data on which actions could be taken for quality improvement (QI). Refer to the following, related report for detailed information on the health plan interviews: “The Utility of CAHPS for Health Plans,” Denise D. Quigley, Dennis Scanlon, Donna O. Farley, and Han de Vries, Santa Monica, Calif.: RAND Corporation, WR-115-AHRQ, 2003. In response to this issue of usability, topic areas that were of high priority to health plans were identified, and items for those topic areas were developed and field-tested.

This report describes the full process undertaken to identify issues, set priorities, and develop and field-test new supplemental QI items for the CAHPS 3.0 health plan survey. The item set is currently being updated for use in the CAHPS 4.0 health plan survey. This information is intended for use by the CAHPS consortium, survey sponsors, health plans, and other stakeholders. Others also may find the information useful as they integrate CAHPS health plan survey items with their own consumer-survey health plan items.

This work was sponsored by the Agency for Healthcare Research and Quality. Charles Darby, Christine Crofton, Kathy Crosson, and Marybeth Farquhar served as project officers at the time of this study.

## **EXECUTIVE SUMMARY**

The Consumer Assessment of Healthcare Providers and Systems (CAHPS®) study was initiated by the Agency for Healthcare Research and Quality (AHRQ) in 1995 to establish survey and reporting products that provide consumers with information on the performance of health plans and providers, as judged by other consumers who have used the health plans. The primary focus of the first CAHPS project was on providing information to help consumers make informed health plan choices.

As use of the CAHPS health plan survey increased, health plans began to express concerns that the survey does not provide them with information on actions that can be taken to help them improve performance in areas important to consumers, as reflected in the survey results. AHRQ initiated a second CAHPS project in 2002 (CAHPS II), working with a consortium of the RAND Corporation, American Institutes for Research, the Harvard Medical School, and Westat. One of the goals of the CAHPS II work was to address the concerns expressed about the need for more action-oriented quality improvement (QI) data from CAHPS.

During CAHPS II, RAND partnered with the National Committee on Quality Assurance (NCQA) and Blue Cross Blue Shield Association (BCBSA) to study the use of CAHPS in health plan QI activities and to develop a set of CAHPS supplemental QI items that health plans can use to collect information for quality improvement activities. This report documents the process undertaken for developing the set of CAHPS supplemental QI items: diagnosing the problem by collecting information from health plans on developmental needs, identifying high priority topic areas for supplemental QI items, developing survey items, and field-testing the items to assess their validity and performance. Each of these steps is detailed in the following paragraphs.

### **DIAGNOSING THE PROBLEM**

To inform efforts to maximize the usefulness of CAHPS for health plan quality improvement, our first step was to diagnose the problem by conducting semi-structured interviews with health plan representatives in a purposive sample of 27 health plans, to gather in-depth information on their quality improvement activities and their perspectives about how CAHPS fits into those activities. In particular, we wanted to develop a better understanding of the priorities and issues of this stakeholder group, which would help us design a CAHPS survey that would indicate actions to be taken and a set of tools for quality improvement.

Because we worked with a purposive sample of plans, the findings from the interviews are not generalizable to all health plans in the United States. However, the interviews identified a strong consensus among these health plan representatives regarding CAHPS and quality improvement, which allowed us to use the information from the interviews with confidence to guide our item development work. Our full findings from the health plan interviews are presented in “The Utility of CAHPS for Health Plans” (Quigley et al., 2003).

Overall, the interviewed health plans indicated that CAHPS has various specific strengths that they value, as well as weaknesses that limit its ability to support their quality improvement activities. The health plans reported that CAHPS was a good, general tool that they can use to compare themselves to other health plans and to examine their performance over time. Its usefulness stems primarily from its scientific soundness, its credibility, and the wide range of topics covered in

the survey. Health plans rated CAHPS's standardization, its capacity for tracking performance over time, and the health plan as its unit of analysis as its best features.

Health plans further reported that CAHPS is limited in its ability to provide information to guide specific actions and interventions in improving their health plan's quality and performance. Reasons cited were that the CAHPS data are reported at the plan level rather than at the provider or medical group level, the content of questions is too general, and results are not relayed quickly enough back to health plans to allow for timely improvements and monitoring of interventions.

Plans do not appear to want to change the existing content areas of the survey, but they are interested in refining items and adding items. In general, the health plans supported inclusion of supplemental items on the CAHPS health plan survey for quality improvement efforts which 1) address issues that a majority of health plans deal with, such as customer service, 2) would assist them in uncovering and focusing on areas in which to take specific action and 3) would help target interventions for improvements in those areas. However, a number of plans voiced concerns about how these common QI issues would be determined, and some also were worried about adding to respondent burden by adding additional items to the survey. Overall, health plans did not want any of the changes to increase costs or survey length, or to result in weakening the credibility, standardization, and comparability that are currently the strong features of the CAHPS health plan survey data.

## **IDENTIFYING PRIORITY TOPIC AREAS**

With the interview feedback from health plans confirming that additional QI supplemental survey questions would be useful, our task was to select topic areas (such as access, customer service, information and education) and specific questions that would be of greatest use to the largest number of plans. We designed a two-step consensus process to accomplish this task, beginning with a Delphi process to identify priority topic areas, followed by review panel sessions to identify specific sub-topics within the priority topic areas for the development of specific survey questions.

The purpose of a Delphi process is to build consensus on a topic among experts in that area, through use of an open forum and structured rating process. All the expert participants in a Delphi process respond individually to sets of questions or rating exercises, and the results of their responses are summarized and fed back to the group in an iterative process (Linstone, 1975; Sackman, 1974).

In this case, we wanted to understand which topic areas are the most important performance areas for ensuring that health plans are serving their enrollees well. The highest priority topic areas are those for which plans are most likely to want good survey information to support their quality improvement work, such as problems underlying why a patient cannot get a referral, or how long it takes for resolution of a customer service issue, or whether office staff were courteous, or whether written materials were easy to understand. We were also interested in understanding how the topic areas might or might not differ in importance if the consumer survey data are reported to the health plan at the individual clinician level, at the medical group level, and/or at the health plan level.

Two rounds of the Delphi process were conducted. In each round, participants performed two rankings of the nine topic areas: access to care, availability of providers, complaints and appeals, provider communication, coordination of care, health plan services, health plan authorization of care, ancillary services, and preventive care. The first ranking was by absolute level of overall importance (ratings), and the second ranking by the importance of each topic area compared with other topic

areas (relative ranking). Of the 34 individuals from 25 health plans invited to perform the topic rankings, 18 individuals from 13 health plans completed the rankings in both rounds.

The three topic areas identified by the Delphi participants as most important at the plan level were

- Health plan services
- Availability of provider
- Access to care.

These same three topic areas were identified as top priorities for the individual clinician survey level. Two of the topic areas—access to care and availability of providers—also were identified as priorities at the medical group level, as was provider communication.

In addition, the ratings and relative rankings for all three levels were analyzed to determine their reliability from the beginning of the Delphi process at Time 1 to the end of the Delphi at Time 2, representing the stability of answers given by the experts between the two rounds of the Delphi, and the consistency between ratings and rankings at each round indicating the level of agreement concerning the importance of a particular topic at a given time period. The estimated reliability of ratings at the plan level was  $r = 0.79$  (intraclass correlation = 0.18) at Time 1 and  $r = 0.92$  (intraclass correlation = 0.42) at Time 2. Thus, the reliability of the ratings for the plan level increased over time.

To examine the level of importance of a particular topic at a given time or across the two time periods, we examine the consistency and stability of the importance ratings and relative rankings by estimating four different sets of correlations for the importance ratings and relative rankings of topic areas. We found that the relative rankings and the importance ratings at all three levels were consistent at the beginning of the Delphi at Time 1. At the second round of the Delphi at Time 2, the relative rankings and importance ratings for the health plan level were consistent only for access, complaints and appeals, and health plan services, indicating that the importance ratings and rankings were providing unique and different information about the importance of the topic areas. Moreover, the relative rankings were consistent during the course of the Delphi process, i.e., from Time 1 to Time 2, at the health plan level for all of the topic areas except availability of providers. Most important, the relative rankings of the topic areas that health plan representatives ranked of highest importance by the Delphi process--access, coordination of care, and health plan services--were significantly correlated at the beginning and the end of the Delphi, i.e. from Time 1 to Time 2, at the health plan level indicating that there was a high level of agreement on the overall importance of these topic areas for quality improvement work by the experts at the beginning and the end of the Delphi process.

In particular, two of the topic areas that were ranked of highest importance—access and health plan services—had ratings and relative rankings that were significantly correlated at the health plan level at the end of the Delphi at Time 2 (indicating consensus between the measures), as well as relative rankings that were correlated from the beginning to the end of the Delphi process from Time 1 to Time 2. In addition, the health plan services topic was significantly correlated at the plan level across all four sets of correlations. The access to care topic area had significant correlations across time for its relative rankings (but not for its ratings), as well as between the ratings and relative rankings at Time 2. The provider availability topic area, on the other hand, only had relative rankings and ratings that were correlated at Time 1, but not in Time 2 or across time indicating that the experts agreed that provider-availability was an important topic area during the first round of the

Delphi, but that by the second round of the Delphi they had reduced its level of importance overall and compared to the other topic areas considered for quality improvement work.

These correlations validated the overall rankings identified by the Delphi participants for the top two topic areas—access to care and health plan services. Given that the health plan representatives ranked the availability of provider topic highly across all three of the organizational levels, we kept it as a high priority topic area, even over provider communication.

## **DEVELOPING SPECIFIC ITEMS WITHIN THE PRIORITY TOPIC AREAS**

To gain feedback and guidance from health plans regarding survey items within the priority topic areas identified in the Delphi process, we held three different review panel meetings with health plans that had participated in both rounds of the Delphi process. The members of the first panel represented health plans that were sophisticated users of consumer-reported data and that were supplementing CAHPS data with QI data from other sources. The members of the second review panel were from health plans that relied more heavily on CAHPS than did the first review panel, although some of them also used other data sources in addition to CAHPS. The third review panel members were from health plans that used only CAHPS as their survey data source.

Based on the priority topic areas identified in the Delphi process, the following topic areas were selected and presented to the review panel for the development of new survey items:

1. Access to care (which we broadened to include availability of providers)
2. Coordination of care (which we added because it has been an area for which developing reliable and useful survey items has been a challenge)
3. Complaints and appeals (component of health plan services)
4. Customer service (component of health plan services)
5. Claims and paperwork (component of health plan services)
6. Information and materials (component of health plan services)

Overall, the three review panels achieved consensus and identified four priority topic areas—access to care, coordination of care, information and materials about health plan services, and customer service. Despite some differing opinions regarding the order of specific sub-topics within topic areas, the three distinct types of health plans represented on the panels agreed upon which topic areas were important for development of supplemental items for health plan quality improvement work. In addition, the three panels provided valuable feedback and guidance about the actual wording and content of specific survey items.

## **FIELD-TESTING CANDIDATE SUPPLEMENTAL QI ITEMS FOR CAHPS**

The culmination of the development process described in this report was the crafting of a set of supplemental QI items for the CAHPS health plan survey. Using the guidance from the health plans described above, we established a set of supplemental QI items covering the following topic areas and sub-topics:

1. Coordination of care
  - a. Communication across providers
2. Access to care
  - a. Appointments for routine care

- b. Appointments with specialists
  - c. After-hours calls
- 3. Information and materials
  - a. Effectiveness of information provided to consumer
  - b. Usefulness of online information and services
- 4. Customer service
  - a. Resolution of issues by telephone customer service
  - c. Representative's knowledge and effectiveness
  - b. Representative's politeness and giving of his/her name.

A total of 23 supplemental QI items were developed, and two modifications for current CAHPS items were suggested. To identify candidate items, we searched the RAND CAHPS project's internal, extensive archive of survey items that has developed over the last 7 years of CAHPS survey development work, and we also wrote some items specifically for this purpose. In addition, several of the health plans that participated in the review panels sent us items covering issues being addressed by their own supplemental QI items, such as, the main reason you had a problem getting a referral to a specialist, or if your customer service issue was not resolved on the first call, how many calls did it take to resolve the issue? Candidate items were refined by the RAND Survey Research Group to conform as closely as possible to CAHPS standards and a final set of items was selected. In some cases, items were allowed to differ from the CAHPS format when it was necessary to preserve their ability to collect the data for action desired by the health plans.

Because the health plans from which we obtained feedback varied widely in their information needs and ratings of importance of items, we took an inclusive approach to identification of survey items, developing a menu of supplemental items from which health plans could choose to fit their specific needs.

Under the standards applied to all CAHPS survey items, candidate items for the supplemental QI item set not only had to address priority topic areas for the health plans but also had to perform well psychometrically. To test the psychometric performance of the candidate QI items we had selected, we partnered with three health plans to add these items to the CAHPS 3.0 health plan survey they fielded in their 2004 annual survey process. Two of the health plans fielded all of the candidate items, and one plan fielded a subset of items that were most relevant to its needs.

Our analyses of the survey data for the supplemental items addressed the following four questions:

- How well do the supplemental QI items perform psychometrically?
- How well do the supplemental QI items discriminate performance across health plans?
- How unique or similar are the QI items to any one of the summarized CAHPS scores for a given topic area, i.e. a CAHPS composite<sup>1</sup>?

---

<sup>1</sup> For example, in the Health Plan level CAHPS survey there are composites in the areas of Getting needed care (4 items), Getting care quickly (4 items), Doctor communication (4 items), Courtesy and respect of office staff (2 items), and Health plan customer service, information and paperwork (3 items).



- Do each of the QI items within a topic area add unique QI information or are there redundancies among the supplemental items? Does the plan have a choice between QI items within a given topic area?

The findings and guidelines for use for these four questions are delineated in the following paragraphs.

### **Psychometric Performance of Supplemental QI Items**

Most proposed supplemental QI items and composites were found to have good psychometric properties. Exceptions were a few items that were applicable to very few people (e.g., those needing after-hours care) or that had more responses in the “other reason” response option than would be desirable. The responses to the supplemental QI items are well distributed. In terms of individual items, the “customer service” item, “In the last 12 months, how often did our health plan’s customer service staff treat you with courtesy and respect?” (Q41A) could possibly have a ceiling effect, which means that the data is consistently near the highest value on the scale (i.e. at the “ceiling”) and therefore the responses for the item have little to no variation, but, given the context of the question, this possible effect should not be a problem.

There were very few responses to the “after-hours care” questions, which read

- QA: After-hours care is health care when your usual doctor’s office or clinic is closed. In the last 12 months, did you need to visit a doctor’s office or clinic for after-hours care?
- QB: In the last 12 months, how much of a problem, if any, was it to get the after-hours care you needed?
- QC: What is the main reason you had a problem getting after-hours care?
- Q5C: What was the main reason you had a problem talking to your personal doctor or nurse by telephone after regular office hours?

Less than 10 percent of the respondents in our sample answered these after-hours items, which indicates that, unless a health plan is very interested in information on after-hours care for its population, it may not a very useful set of items because few people indicated on the survey that they seek access to after-hours care.

Finally, three supplemental QI questions that ask about reasons for different problems have large percentages of responses in the “other reason” response option. Additional field-test research is scheduled for these three items to identify which additional response categories are needed, and they will be revised based on the information being gathered. These items are as follows:

- Q9A: What was the main reason you had a problem seeing a specialist?
- Q5C: What was the main reason you had a problem talking to your personal doctor or nurse by telephone after regular office hours?
- Q42A: What was the main reason you had a problem getting help from your health plan’s customer service?

**Guidelines for use:** All of the supplemental QI items for “getting care quickly,” “health plan customer service” (except Q42A, which may need more response options), and “coordination of care” can be used with confidence about their psychometric properties. The majority of the supplemental items for the topic of “getting needed care” also had good psychometric properties.

However, very few patients responded to the items concerning “after-hours care” (QB, QC, and Q5C), indicating that these items may not be useful overall to a health plan because very few people need after-hours care.

### **Supplemental QI Items’ Ability to Discriminate Performance Across Health Plans**

Regression models were run; each with a QI item as the dependent variable and the health plans and plan products as the independent variables. The responses to the supplemental QI items varied by plan product, with the exception of a few items that were compared across only a few products. For all but four items, the reliability measures were high and in the reliable range, showing that the items discriminated performance across health plan products. We also conducted these analyses using only the health plan as the independent variable, but with only three plans field-testing the items, there were not enough plans to generate definitive results at the plan level.

**Guidelines for use:** The supplemental QI items are able to discriminate variation in performance across products and have high reliability, except questions QB, Q41A, Q41B, and Q34E:

- QB: In the last 12 months, how much of a problem, if any, was it to get the after-hours care you needed?
- Q41A: In the last 12 months, how often did your health plan’s customer service staff treat you with courtesy and respect?
- Q41B: In the last 12 months, how often did you hang up the phone feeling certain your health plan’s customer service staff would get you the help you needed?
- Q34E: How satisfied are you with the help you received to coordinate your care in the last 12 months?

### **Similarity of the Supplemental QI Items to the CAHPS Composites**

The majority of the candidate items had moderate correlations with their intended composites, and thus should provide useful information for quality improvement efforts. The only two exceptions were Q5A, which has a low correlation with the “getting needed care” composite (0.04) and Q40C and Q40E, which have low correlations with the “health plan customer service” composite. These items read as follows:

- Q5A: In the last 12 months, did you call your personal doctor or nurse during regular office hours?
- Q40C: When you looked for information in the last 12 months, did you go to your health plan’s Internet site?
- Q40E: In the last 12 months, did you use information on your health plan’s Internet site to help choose a health care provider?

The QI items for “coordination of care” are not correlated as a group to any one CAHPS composite. However, individual “coordination of care” items are correlated with the “getting needed care” composite (Q34B and Q34E) and “getting care quickly” composite (Q34B). Specifically, these “coordination of care” items read:

- Q34B: In the last 12 months, how often did your personal doctor or nurse seem informed and up-to-date about the care you received from specialists or different kinds of health care providers?
- Q34E: How satisfied are you with the help you received to coordinate your care in the last 12 months?

**Guidelines for use:** With moderate correlations between most of the QI items and their intended CAHPS composites, the QI items should be useful to health plans in focusing their interventions on actions that can improve their overall scores on those composites. In addition, given the cross correlations between performance on “after-hours care” and “customer service” items, improvements made for one of these areas should also help improve performance in the other area. Moreover, only three QI items- (Q5A) Call personal doctor after hours, (Q40C) Look for information on plan Internet site, and (Q40E) Use plan Internet site information to choose a doctor--have very low correlations with their relevant composite; therefore, it would not be advisable for health plans to use them for improving scores on that composite. However, these three items could be useful to a health plan that is specifically interested in making improvements in the area measured by the specific QI item. Finally, the QI items and the composites are not highly correlated, indicating that they do not reflect the same content. Therefore, a health plan would want to include both the composite score and the QI item because they each have unique information to provide to the health plan.

### **Uniqueness or Redundancy of Supplemental QI Item Information Within a Topic Area**

The supplemental QI items are correlated with each other, and the correlations are moderate (i.e., above .4 to .5), indicating that there are no redundancies among the items and that each item has unique information to provide. The supplemental QI items were also predictive of their related composites, indicating that the supplemental QI items do add information to the composites, both individually and collectively. Regression analyses were performed that identified the following subset of *a priori* hypothesized supplemental QI items that provide the most additional information for a given composite with the least number of additional QI items needed:

- The item Q9A, “What was the main reason you had a problem seeing a specialist?” was moderately predictive ( $R^2$  of 0.23) of the “getting needed care” composite.
- Three items were highly predictive ( $R^2$  of 0.60) of the “health plan customer service” composite. These were Q42B, “How many calls did it take for you to get the help you needed from your health plan’s customer service?” Q42A “What was the main reason you had a problem getting help from your health plan’s customer service?” and Q40A, “What kind of information did you have a problem understanding or using?” (Specifically, the choice--getting care outside of your network).

**Guidelines for use:** All of the supplemental QI items with moderate correlations to their related composites could generate usable information for QI if they were included on the survey. At the same time, a smaller number of key items have the strongest influence on each composite, and these items could be selected for inclusion in a survey for which there is space for only a few supplemental items.

For example, if a health plan is interested in improving its overall composite score on “getting needed care,” then it could gather data on the “access to needed care” supplemental QI items, to assess how well the plan is performing on each supplemental QI item. If performance on

one or more item is lower than desired, then the plan could work to improve performance on those QI items with reasonable confidence that improvements in those areas would help increase its overall score on the “getting needed care” composite. If the health plan could include only one QI item on its survey, and wanted to have the greatest impact on the composite, then it should use the ones with the stronger correlations to the composites. In terms of “getting needed care,” this one supplemental QI item would be Q9A, “What was the main reason you had a problem seeing a specialist?”

## **CONCLUSION**

We anticipate that the supplemental QI items developed through this extensive process will be useful to health plans in gathering information for taking actions to better focus their QI interventions to improve their overall CAHPS composite scores. It will be important to track the extent to which health plans choose to use these measures, as well as to continue to explore additional topic areas that may also be of importance to them, to ensure that the item development process is responsive to their needs and priorities. We also anticipate that there will be demand for similar sets of supplemental QI items related to other CAHPS surveys.

