



HEALTH

- THE ARTS
- CHILD POLICY
- CIVIL JUSTICE
- EDUCATION
- ENERGY AND ENVIRONMENT
- HEALTH AND HEALTH CARE
- INTERNATIONAL AFFAIRS
- NATIONAL SECURITY
- POPULATION AND AGING
- PUBLIC SAFETY
- SCIENCE AND TECHNOLOGY
- SUBSTANCE ABUSE
- TERRORISM AND HOMELAND SECURITY
- TRANSPORTATION AND INFRASTRUCTURE
- WORKFORCE AND WORKPLACE

This PDF document was made available from www.rand.org as a public service of the RAND Corporation.

[Jump down to document](#) ▼

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.

Support RAND

[Purchase this document](#)

[Browse Books & Publications](#)

[Make a charitable contribution](#)

For More Information

Visit RAND at www.rand.org

Explore [RAND Health](#)

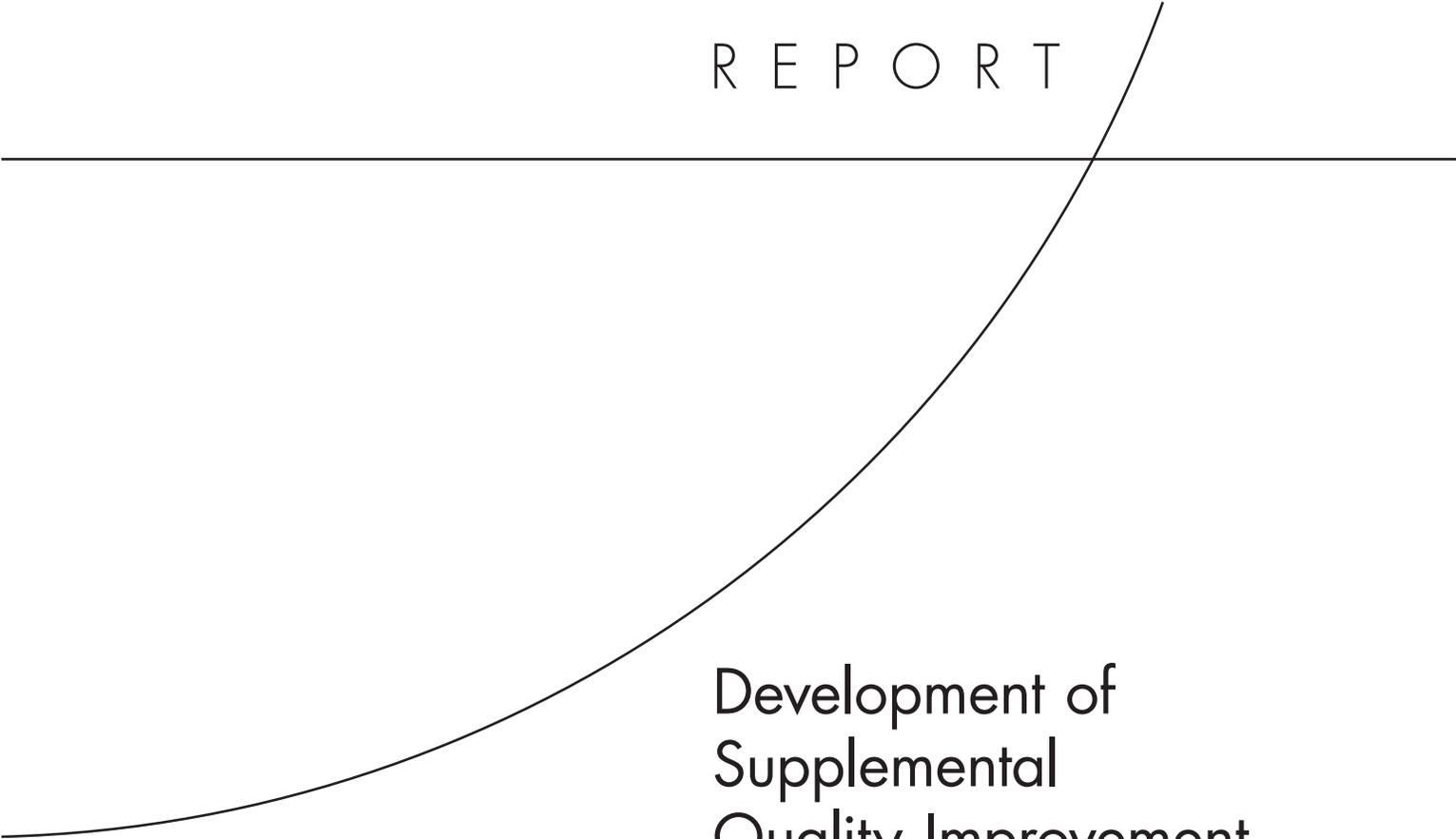
View [document details](#)

Limited Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law as indicated in a notice appearing later in this work. This electronic representation of RAND intellectual property is provided for non-commercial use only. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use.

This product is part of the RAND Corporation technical report series. Reports may include research findings on a specific topic that is limited in scope; present discussions of the methodology employed in research; provide literature reviews, survey instruments, modeling exercises, guidelines for practitioners and research professionals, and supporting documentation; or deliver preliminary findings. All RAND reports undergo rigorous peer review to ensure that they meet high standards for research quality and objectivity.

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.

© 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

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.

TABLE OF CONTENTS

Preface	iii
Tables	vii
Executive Summary	ix
Acknowledgments	xix
CHAPTER ONE THE CONTEXT: MAKING CAHPS MORE USABLE FOR QUALITY IMPROVEMENT	1
CHAPTER TWO DIAGNOSING THE PROBLEM	3
Health Plan Use of Consumer-Reported Information.....	4
Use of Consumer-Reported Information in General	4
Health Plans’ Use of CAHPS	5
Feedback on CAHPS’ Strengths and Weaknesses.....	6
Summary.....	7
CHAPTER THREE IDENTIFYING PRIORITY TOPIC AREAS	9
Design of the Delphi Process.....	9
Results from the Delphi Process.....	11
Summary.....	13
CHAPTER FOUR DEVELOPING SPECIFIC ITEMS WITHIN THE PRIORITY TOPIC AREAS	15
The Review Panel Process.....	15
Results of the Three Review Panels.....	17
The First Technical Review Panel.....	17
The Second Technical Review Panel	19
The Third Technical Review Panel	21
CHAPTER FIVE TESTING CANDIDATE SUPPLEMENTAL QI ITEMS FOR CAHPS	23
Field-Testing of the Supplemental QI items.....	24
Field-Test Results	24
1. How well do the supplemental QI items perform psychometrically?	25

2. How well do the supplemental QI items discriminate performance across health plans? ...	28
3. How unique or similar are the QI items to the CAHPS composites?	30
4. Does each of the QI items within a topic add unique QI information, or are there redundancies among the supplemental items? Does the plan have a choice between QI items within a topic?.....	32
CHAPTER SIX DISCUSSION AND CONCLUSION	37
Identification of Important Issues for Health Plans	37
Usability of the Supplemental QI Items.....	38
Conclusion	38
APPENDIX A DESIGN AND METHODS FOR THE HEALTH PLAN INTERVIEWS	39
APPENDIX B RELIABILITY OF RATINGS OF IMPORTANCE OF QI ITEMS.....	45
APPENDIX C NEW CAHPS SUPPLEMENTAL ITEMS FOR QUALITY IMPROVEMENT	51
APPENDIX D SUPPLEMENTAL QUALITY IMPROVEMENT (QI) ITEMS IN THE FIELD TEST	53
APPENDIX E CORRELATIONS OF SUPPLEMENTAL QI ITEMS	61
REFERENCES.....	63

TABLES

Table 3.1	Nine Topic Areas Rated in the Delphi Process.....	10
Table 3.2	The Top Three Most Important Topic Areas Identified by Delphi Participants, by Organizational Level of Representative.....	11
Table 3.3	Correlations Between Importance Ratings and Relative Rankings of Topic Areas from the Two Rounds of the Delphi Process.....	12
Table 4.1	Ranking of Topic Areas and Sub-topics by the First Technical Review Panel.....	18
Table 4.2	Rankings of Topic Areas and Sub-topics by the Second Technical Review Panel.....	20
Table 5.1	Psychometric Properties of QI Items, Grouped by Topic Area and Composite.....	27
Table 5.2	Regression Results with QI Items as the Dependent Variables and Plans with Products as Independent Variables.....	29
Table 5.3	Correlations of CAHPS Composites and Quality Improvement Items.....	31
Table 5.4	Regression Results, with Composites as Dependent Variables and the <i>A Priori</i> Corresponding Set of Quality Improvement Items as Independent Variables.....	35
Table 5.5	Regression Results for Composites as Dependent Variables, Using Stepwise Regression to Include QI Items in Blocks as Independent Variables.....	36
Table A.1	Distribution of People Interviewed, by Type of Position Held.....	42
Table A.2	Total Enrollment Size of the Health Plans Included in Interviews.....	43
Table A.3	Distribution of Model Types for the Health Plans Included in Interviews.....	43
Table B.1	Partitioning Between-Topic-Area Variance and Within-Topic-Area Variance at Time 1.....	47
Table B.2	Partitioning Between-Topic-Area Variance and Within-Topic-Area Variance at Time 2.....	47
Table B.3	Spearman Correlations Between Ranks and Ratings of Importance of Plan Topic Areas at Time 1 ($N = 18$).....	48
Table B.4	Spearman Correlations Between Ranks and Ratings of Importance of Plan Topic Areas at Time 2 ($N = 17$).....	48
Table B.5	Spearman Correlations Between Ranks of Plan Topic Areas from Time 1 to Time 2 ($N = 17$).....	49
Table B.6	Spearman Correlations Between Ratings of Plan Topic Areas from Time 1 to Time 2 ($N = 17$).....	49
Table E.1	Correlations of Supplemental QI Items.....	61

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 composite¹?

¹ 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.

ACKNOWLEDGMENTS

We are indebted to the health plans and health plan representatives whom we interviewed and who participated in the Delphi process and the review panels. They were generous with their time and participated thoughtfully in the process, providing us with invaluable feedback and guidance for identification of priority topic areas and supplemental items to help support health plans' quality improvement activities. We are also very grateful to the National Committee on Quality Assurance and the Blue Cross Blue Shield Association of America for their support and participation in identifying issues to be considered and for serving as liaisons with the health plans as we recruited plans for interviews and carried out the subsequent item selection activities.

A number of other individuals made important contributions to the development of this work. For their valuable input, we thank Marla Haims, Harry Holt, Marika Suttrop, and Carol Edwards, all of who are colleagues on the RAND CAHPS team

CHAPTER ONE

THE CONTEXT: MAKING CAHPS MORE USABLE FOR QUALITY IMPROVEMENT

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 health plan performance, as judged by other consumers who have used the health plans. For the initial development and refinement of the CAHPS products, AHRQ worked with a research consortium consisting of the RAND Corporation, Research Triangle Institute, Harvard Medical School, and Westat, an employee-owned contract research organization founded in 1961 headquartered in the Washington, DC area. The primary focus of CAHPS was on providing information to help consumers make informed health plan choices.

The CAHPS health plan survey measures health plan performance on several dimensions, including global ratings (e.g., ratings of health plan, primary doctors or nurse, specialty care) and reports of experiences with using a health plan (e.g., doctor or nurse listened and showed respect, could get an appointment when needed, got specialty referral when needed, plan responded to concerns or questions). The ratings are individual items using response scales ranging from 0 to 10. The reports of experiences are composite scores, which are averages of responses to sets of individual items with four response options (Agency for Healthcare Research and Quality, 1999).

It was learned during the first CAHPS development work that various stakeholders have very different perspectives and uses for CAHPS. (Carman et al., 1999; Fox et al., 2001; Spranca et al., 2000; Kanouse et al., 2000; Farley et al., 2002a, 2002b). These stakeholders include not only consumers but also sponsors, health plans, and providers that use CAHPS.

As use of the CAHPS health plan survey increased, health plans began to express concerns that the survey did not provide them with information on actions they could take that could help them improve performance in areas important to consumers, as reflected in the survey results. Both the National Committee on Quality Assurance (NCQA) and the Blue Cross Blue Shield Association (BCBSA) had received feedback from their constituent health plans regarding these same concerns.

AHRQ initiated a second CAHPS project in 2002 (CAHPS II), working with a consortium of the RAND Corporation, American Institutes for Research, Harvard Medical School, and Westat. One of the goals of the second CAHPS work was to address the concerns expressed about the need for more data from CAHPS on actions that could be taken.

During CAHPS II, RAND has pursued a coordinated strategy to strengthen the usability of CAHPS for health plan quality improvement actions. As part of this strategy, we partnered with NCQA and BCBSA to study the use of CAHPS in health plan quality improvement activities and to develop a set of CAHPS supplemental QI items that health plans can use to collect information on such actions for quality improvement. A draft set of QI items was developed in 2003, using a participative process with health plans, the BCBSA, and NCQA. These items were field-tested in 2004 in partnership with three health plans.

This report documents the process undertaken for developing the set of CAHPS supplemental QI items. The four steps in the process were (1) diagnosing the problem by gathering information via interviews from health plan representatives on their survey development and quality improvement needs, (2) identifying priority topic areas for supplemental QI items, (3) identifying

issues with any of the priority topic areas or sub-topics, and, finally, (4) developing survey items and field-testing them to assess their validity and performance. Each of the chapters of this report addresses one of these steps and discusses what we learned in each step.

This report summarizes the findings and lessons learned from RAND's development of CAHPS supplemental items for quality improvement. In this chapter, we describe the context for and background on the need for making CAHPS more usable for quality improvement. Chapter Two presents the results of our interviews with health plan representatives, through which we learned their perspectives about how CAHPS fits into their QI process and the limitations of CAHPS data for use in QI. Chapter Three describes the Delphi process we facilitated, through which participating health plans identified topic areas that were important to them for quality performance, and Chapter Four describes the review panel process used to obtain health plan guidance on identifying important sub-topics within those topic areas. Chapter Five describes how we developed the supplemental QI items for the topic areas identified by the health plans, and it presents results from our analysis of the psychometric performance of the CAHPS QI supplemental items, using field-test data for the items. Finally, Chapter Six summarizes the key findings from this work and the implications of those findings for health plans that may choose to use the supplemental QI items.

CHAPTER TWO

DIAGNOSING THE PROBLEM

To inform efforts to maximize the usefulness of CAHPS for health plan quality improvement, our first step was to conduct semi-structured interviews with health plan representatives in a purposive sample of health plans in three categories: 1) those health plans that do not field CAHPS, 2) those that field CAHPS but do not publicly report their CAHPS results in the NCQA's Quality Compass database¹, which is database of health plans that allows NCQA to release their results and is used as a tool for selecting a health plan, conducting competitor analysis, examining quality improvement, and benchmarking health plan performance, and 3) health plans that performed across the spectrum on CAHPS--from consistently higher than average to those consistently lower than average. Our goal was 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 with items on which action could be taken and a set of tools for quality improvement. These interviews were a component of a broader set of market research interviews conducted by the CAHPS consortium during 2003.

Some work had been undertaken by RAND and Harvard to document how health plans have used CAHPS in quality improvement activities. However, health plans' use of CAHPS is multifaceted, and we are only beginning to understand the patterns of use and related issues. Thus, we needed to use a data collection methodology that offered the flexibility for respondents to share their unique experiences with us. We also wanted to obtain structured feedback on specific design changes for the CAHPS health plan survey, along with open-ended responses on the utility of each change and suggestions for design. Our purposive sample contained 4 health plans that did not use CAHPS; 4 health plans that fielded CAHPS but did not report results publicly, and 4 or 5 plans in 4 of the performance categories: consistent high performers, consistent low performers, improvers, and those who declined in performance. We did not sample any health plans in the "Mixed" performance category. These 27 health plans were distributed across the regions of the United States: East ($N=10$), Midwest ($N=7$), North ($N=5$), West ($N=4$) and South ($N=1$). The majority of our sample ($N=11$) had a mixed structure combining a network model and an independent physician association (IPA); about one-third of our sample ($N=8$) had a network model; and approximately 20 percent of the sample ($N=5$) had independent physician association (IPA) models. Refer to Appendix A for a description of the design and methods for the health plan interviews, as well as for more details on the characteristics of the sample.

This design allowed us to explore the usefulness of CAHPS with health plan representatives who had considerable experience in performance measurement for quality improvement. Moreover, the interviews identified a strong consensus among these health plan representatives regarding CAHPS and quality improvement. Because of this consensus, we were able to use the information from the interviews to guide our item development work. We would have had less confidence in applying the results if there had been a greater diversity of opinions among the respondents.

The details of our findings from the health plan interviews are presented in "The Utility of CAHPS for Health Plans" (Quigley et al., 2003). Highlights of our findings are presented in this chapter.

¹ For more information, please refer to: www.ncqa.org/Info/QualityCompass/index.htm.

Overall, the interviewed health plans indicated that CAHPS has various specific strengths that they value, as well as weaknesses that limit their ability to support their quality improvement activities. The health plans found CAHPS to be a good, general tool that they can use to compare themselves to other health plans and to track performance over time and benchmarking. Its usefulness stems primarily from the credibility that the data has among the health care community because of its scientific reliability and data collection standards as well as the array of topics covered in the survey. Plans do not appear to want to change the existing content areas of the survey, but they are interested in refining items and in adding items.

HEALTH PLAN USE OF CONSUMER-REPORTED INFORMATION

To best understand how health plans use CAHPS as an information source, it is important to have information on the priority placed by plans on consumer-reported information compared with other types of information on their performance. Information was gathered regarding which departments within the health plans used consumer-reported information, and their views and use of CAHPS as an information source. Each subsection below highlights our main findings.

Use of Consumer-Reported Information in General

A majority of the health plans interviewed felt strongly that reports by consumers on their experience with care and customer service were equally important or more important than other performance issues, such as clinical quality measures or cost information. The majority of health plan representatives also indicated that their organizations placed a high priority on measuring and improving performance according to consumer-reported measures.

Highlights:

- 41 percent of the health plans indicated that consumer reports on their health plan experiences were more important than other performance issues, such as clinical quality or cost, and another 52 percent indicated that they were equally as important as those other issues.
- The main reasons why consumer-reported measures were important to health plans were expectations by customers for effective service, competitive pressures to be service-driven or customer-focused, and the nature of consumer-reported measures as the only way to understand what consumers think and how they interact with their health plans.

Use of consumer-reported measures is widely distributed throughout the health plans. The demand for consumer-assessment information primarily comes from quality improvement (QI)-related departments, but also from senior management, marketing, and departments overseeing accreditation preparation.

Highlights:

- 70 percent of the health plans interviewed indicated that consumer-reported data, including CAHPS and other types of data, were sent to quality improvement departments, committees, and teams
- The demand for consumer-reported data within the health plans came from various and multiple sources: QI departments, committees and teams (44 percent of plans), senior management (32 percent), marketing (28 percent), and accreditation (28 percent).

In performance reports using consumer-reported measures, health plans indicated that benchmarks and trend analysis were the two most important elements of these reports, and regional

data were more useful to their organization when making comparisons or benchmarking than were local or national data.

Highlights:

- 89 percent of the health plans indicated that benchmarks were important in reporting consumer-reported measures, and 70 percent also indicated trend analysis was important.
- 78 percent of the health plans compare consumer-reported measures with benchmarks.
- 85 percent of the health plans indicated that regional data were the most useful to their organization.

Health Plans' Use of CAHPS

CAHPS is an important source of consumer-reported data that has credibility with the health plans. CAHPS was reported by health plans as credible in terms of its scientific integrity and the comprehensiveness of the topic areas covered in the survey, but a smaller percentage of health plans found CAHPS useful in terms of the specificity of information generated from the survey.

Highlights:

- 76 percent of the health plans found CAHPS credible in terms of scientific integrity.
- 64 percent of the plans found CAHPS credible because of topic areas.
- 40 percent of the plans found it credible because of the specificity of information generated by the survey.

A large majority of health plans use CAHPS consumer-reported data in conjunction with other surveys and data sources. Most often health plans use the adult CAHPS survey and analyze the data only at the plan level. At that time, the CAHPS survey was not available at the group and clinician levels, as is CAHPS 4.0 now.

Highlights:

- 78 percent of health plans interviewed indicated that they use CAHPS in conjunction with other survey data
- All 4 of the health plans that were selected because they “currently did not use CAHPS” indicated two main reasons why they did not use CAHPS: 1) the cost of the data collection and sampling procedures to administer the CAHPS survey is higher than their cost of administering other non-standardized or home-grown consumer-experience surveys and 2) CAHPS did not fulfill all of their specific survey and data needs. [NOTE: NCQA requires that health plans report CAHPS health plan survey scores for accreditation; thus, the health plans perceive that requiring CAHPS creates a “cost” for accreditation both in terms of the higher expenditures to administer the CAHPS survey and in terms of not having all of their data and survey needs met via CAHPS.]

Health plans reported using CAHPS data in quality improvement, credentialing, and marketing. CAHPS is used in their quality improvement efforts because it allows for benchmarking, provides valid and reliable data, and provides consumer-reported data that is not available from another source.

Highlights:

- 84 percent of health plans indicated that CAHPS was useful for quality improvement activities.
- 44 percent of health plans indicated that actions to be taken couldn't be derived from CAHPS data, 30 percent indicated that they can be, and 26 percent indicated that actions to be taken depend on the topic area addressed in the survey.
- 85 percent of the health plans reported that inadequate data specificity was the primary barrier that restricted CAHPS from playing a more integral role in quality improvement.

For quality improvement, CAHPS data are used most often in customer service. After customer service, four areas were of similar priority: complaints and appeals, access to care, availability of providers, and paperwork and claims.

Highlights:

- 25 percent of priority measures reported by health plans were in the area of customer service.
- After customer service, several areas were of equal priority in the CAHPS data: access to care, availability of providers, and paperwork and claims.

FEEDBACK ON CAHPS' STRENGTHS AND WEAKNESSES

CAHPS is used primarily for tracking performance over time, benchmarking, making comparisons with other health plans, and identifying areas for quality improvement. Moreover, health plans rated CAHPS' standardization and capability for tracking performance over time as its best features.

Highlights of Overall Results:

- 85 percent felt health plan representatives considered the main strength of CAHPS to be its usefulness for tracking performance over time and for benchmarking for comparisons with other health plans.
- 78 percent used CAHPS primarily to identify quality improvement issues
- 44 percent used CAHPS because it is required for NCQA accreditation
- 63 percent felt the topics covered by CAHPS were appropriate
- Topic areas on CAHPS that were most useful are:
 - customer service (89 percent)
 - access to care (59 percent)
 - claims and paperwork (59 percent)
- Topic area on CAHPS that was least useful—provider communication (67 percent)
- 30 percent wanted more questions on health plan customer service issues
- 19 percent disliked the negative framing of the “problem” CAHPS questions
- 22 percent felt that the questions on referrals to specialists were too focused on health maintenance organizations (HMOs).

Health plans further reported that CAHPS is limited in its ability to establish 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 only and do not include information at the provider

level; that the contents of questions are too general; and that results are not relayed quickly enough back to health plans to allow for timely improvements and monitoring of interventions.

The content topic areas covered in the CAHPS health plan survey are relevant to health plans, but within a few of the topic areas, health plans want to add more items and make some specific changes to existing items.

Highlights of Results on CAHPS Limitations:

- 74 percent of health plan representatives identified the limited specificity of the survey, including scope of survey items, unit of analysis (plan, group, provider), and type of health plan, as a drawback
- 41 percent identified obtaining the survey data long after the survey was administered as a drawback
- 70 percent wanted provider data aggregated at the group or individual provider level.

Given these limitations, the health plans reported that they tend to supplement CAHPS with other, more real-time data (i.e., claims data and operations data) that are specific to their markets and to the content areas they have identified as areas needing improvement. In addition, by supplementing the CAHPS data with other survey data, the health plans can pinpoint more specific issues on which they need to improve, to set and establish goals, and to monitor these interventions. In this ongoing process of quality improvement, health plans report that CAHPS data assist primarily in identifying the general area(s) for improvement. However, the CAHPS data themselves are not diagnostic enough to identify needed improvements solely on their own. Additional information is needed to define the specific aspects that need to be improved.

SUMMARY

Health plans supported the development of supplemental items on the CAHPS 3.0 health plan survey to enable them to gather information that is useful for their most common quality-improvement efforts. Data from these supplemental items would assist in obtaining focused, specific data on issues where specific actions could be taken for improvement and would help to better define and target interventions for improvement. However, a number of plans voiced concerns about how the common QI issues would be determined. These concerns were echoed by plans that did not support this design change (and by those worried that respondent burden would increase). Overall, health plans did not want changes to increase survey costs or survey length, or to result in weakening of the credibility, standardization, and comparability that are such strong features of the CAHPS survey data.

CHAPTER THREE

IDENTIFYING PRIORITY TOPIC AREAS

Given the feedback from interviews of health plans confirming that QI supplemental survey questions would be useful, we next set out to select topic areas 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: a Delphi process to identify priority topic areas, followed by review panel sessions to identify specific topics for survey questions within the priority topic areas. In this chapter, we describe the Delphi process and discuss the priorities that emerged from the process. The review panel process and results are described in Chapter Four.

To our knowledge, no work to date had been undertaken to document the level of importance of particular topic areas for health plans' QI efforts. Thus, we needed to use a methodology that offered an open, but structured, forum for plan representatives to share their opinions with us and to rank the topic areas. We chose to use a Delphi process with representatives from the 27 health plans that we had interviewed. Through the Delphi process, we sought their participation to reach consensus on which of the CAHPS health plan survey topic areas were the highest priority for developing additional survey items to provide more data for quality improvement actions that needed to be taken.

DESIGN OF THE DELPHI PROCESS

The purpose of a Delphi process is to build consensus on a topic among experts in that topic, through use of an open forum and a structured rating process. In a Delphi process, all the expert participants respond individually to sets of questions or rating exercises, the results of which are summarized and fed back to the group in an iterative process. The process is facilitated by a moderator, who develops the initial materials sent to the experts and then coordinates the summarization and feedback after each iteration (Linstone, 1975; Sackman, 1974).

The Delphi process consists of the following basic steps to investigate and understand experts' beliefs and opinions:

1. A targeted audience of experts is asked a question (or series of questions) around which there is a need to gain consensus. Each expert is asked the same question(s) individually in a written format.
2. All the experts' responses are assembled and summarized to reflect the opinions and answers of the group.
3. The summary of results from the first Delphi round is sent back out to the experts, along with the exact set of questions from the first round. Experts are again asked to provide their feedback, taking into consideration the results from the previous round of responses.
4. This process is repeated two or more times and allows a consensus to be achieved.

In this case, we wanted to understand which *topic areas* are most important to health plans for ensuring that they are performing well in serving their enrollees. The highest priority topic areas are those for which plans are most likely to want good survey information to support their quality improvement work. We were also interested in understanding how the topic areas might or might not differ in importance if the consumer survey data are reported at the individual clinician level, at the group level, and/or at the health plan level.

We assembled information from multiple sources about the topic areas used by health plans in quality improvement activities and consumer health care surveys. We identified nine main topic areas, which are presented in Table 3.1.

Table 3.1
Nine Topic Areas Rated in the Delphi Process

Topic Area	Definition
Access to care	How timely and reachable a doctor is by a patient in terms of intake, setting up appointments, and wait time in the office and in the exam room. It includes 24-hour, 7-day-a-week access.
Availability of provider	The actual choice of a provider by a patient, including changing doctors and the referral process to specialists. <i>Specialists</i> are providers whose practices are limited to treating a specific disease (e.g., oncologists), specific parts of the body (e.g., ear, nose and throat), a specific age group (e.g., pediatrician), or specific procedures (e.g., oral surgery).
Complaints and appeals	All the processes within a health plan pertaining to grievances, complaint resolution, and appeals.
Provider communication	The interaction and direct communication between providers and patients.
Coordination of care	A health plan’s approach to managing medical care from prevention through treatment and recovery. It includes such practices as the management of chronic conditions, such as diabetes or asthma, as well as utilization management.
Health plan services	Services a plan offers to its consumers in addition to medical care, such as customer service, a process for obtaining an ID card, all written information and materials on conditions and treatments, and paperwork and claims.
Health plan authorization of care	The decisions a health plan makes regarding the care it provides to customers, ranging from the drug formulary, to behavioral health access, to pre-authorization of services
Ancillary clinical services	Diagnostic and supportive services, such as radiology, physical therapy, pharmacy, and laboratory work.
Preventive care	Medical and dental services aimed at early detection and intervention, including all standard preventive medical practices, such as immunizations and periodic examinations, but excluding medical education.

In each of the two rounds of the Delphi process (referred as Time 1 and Time 2), participants performed two rankings of the nine topic areas—the first by absolute level of overall importance (ratings from 0 to 10), and the second by the importance of each topic relative to each other (relative rankings). We invited 34 individuals from 25 health plans to perform the rankings. (Two of the health plans that we had interviewed had never fielded CAHPS and were not included in the Delphi.) Of these 34 individuals, 18 individuals from 13 health plans completed the rankings in both rounds-- a 72-percent response rate for health plans and a 53-percent response rate for individuals. The Delphi participants were given the following instructions:

1. Overall Rating of Topic Areas

Instructions: Please put the current CAHPS composites and questions out of your mind. Imagine that you could collect closed-question survey data on the above nine topic areas from consumers of your health plan to inform your quality improvement work and your health plan’s performance. What are the most important data that you need to know from your consumers aggregated at the plan level?, at the group level?, and at the individual clinician level? What data is key in making decisions and performing well as a health plan?

Rating scale: 0 to 10, where 0 = least important and 10 = most important

2. Relative Ranking of Topic Areas

Instructions: Now that you have thought about the overall importance of each of these nine topic areas, we would like you to think about their relative importance to each other. Thinking about how important these topic areas are for a survey that provides you with data aggregated at either the individual clinician level, at the group level, or at the plan level. Please RANK how important each topic area is to include in a consumer survey that will inform your work and your health plan’s performance. Please RANK the nine topic areas for each level from the most to least important.

Rating scale: Number the topic areas 1 through 9 in order of importance, using 1 as most important.

RESULTS FROM THE DELPHI PROCESS

The priority topic areas identified by the health plans in the Delphi process varied slightly by the level from which the representative came (plan, group, provider/clinician). The top-identified topic areas as reflected in absolute ratings according to the level of health care delivery are presented in Table 3.2. Availability of providers and Access to care were among the top three regardless of the level of health care delivery. Provider communication was rated among the top three for group level only, and health plan services was among the top three for both plan and clinician levels.

**Table 3.2
The Top Three Most Important Topic Areas Identified by Delphi Participants, by
Organizational Level of Representative
(18 representatives from 13 health plans)**

Topic Area	Organizational Level		
	Plan	Group	Clinician
Health plan services	1		3
Availability of providers	2	2	1-2
Access to care	3	1	1-2
Provider communication		3	

In addition, the overall ratings and relative rankings for all three levels were analyzed to determine their reliability from the first round of the Delphi (referred to here as Time 1) to the second round of the Delphi (referred to here as Time 2) and the consistency between ratings and rankings at each round. Appendix B presents the detailed results of these analyses. We estimated the reliability of the importance ratings for both Delphi rounds. 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, which means that the reliability of the ratings for the plan level increased over time.

To examine the consistency of the importance ratings and relative rankings, we estimated four different sets of correlations for the importance ratings and relative rankings of topic areas (refer to Table 3.3, columns 1-4). Results for the health plan level are reported in Table 3.3. With few exceptions, the relative rankings and the importance ratings across all of the topic areas for all three levels were consistent (with a correlation above .40) at Time 1, the first round of the Delphi prior to influence from other experts’ opinions (refer to column 1, Table 3.3 for the correlations at the health plan level). The three exceptions at the health plan level were access, complaints and appeals, and coordination of care. At Time 2, the second round of the Delphi after participants had reviewed the opinions of the other experts, 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 experts’ opinion on the importance and relative importance of the topic areas (refer to column 2). As a result, both the importance ratings and ranking were then used to establish the top three topic areas.

Moreover, the relative rankings were consistent during both rounds 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 (refer to column 4, Table 3.3). Most importantly, the relative rankings of two of the topic areas that were ranked of highest importance by the Delphi process by the health plan representatives—access (.80) and health plan services (.72)--were among the topic areas that were significantly correlated during both rounds of the Delphi (i.e., from Time 1 to Time 2) at the health plan level.

Table 3.3
Correlations Between Importance Ratings and Relative Rankings of Topic areas
from the Two Rounds of the Delphi Process

Topic Area	Spearman Correlations ^b at the Health Plan Level			
	Column 1: Ratings and Rel. Rankings at Time 1	Column 2: Ratings and Rel. Rankings at Time 2	Column 3: Ratings at Time 1 and Time 2	Column 4: Relative Rankings at Time 1 and Time 2
Access to care	-0.29	-0.52*	0.36	0.80**
Availability of provider	-0.64**	-0.00	0.28	0.33
Complaints and appeals	-0.37	-0.63**	0.43	0.60**
Provider communication	-0.72**	-0.48	0.84**	0.66**
Coordination of care	-0.19	-0.32	0.43	0.67**
Health plan services	-0.74**	-0.63**	0.83**	0.72**
Health plan authorization of care	-0.82**	-0.20	0.62*	0.55*
Ancillary clinical services	-0.49*	-0.39	0.73**	0.69**
Preventive care	-0.62**	-0.14	0.81**	0.48*

NOTE: The statistically significant results for the topic areas that were ranked of highest importance as seen in Table 3.2 are **bolded** in this table. Statistical significance is denoted by * $p < .05$, ** $p < .01$.

^a These correlations are negative because “most important” is a 10 on the rating scale and a 1 on the relative-ranking scale.

^b Spearman's rank correlation coefficient, often denoted by the Greek letter ρ (rho), is a non-parametric measure of correlation—that is, it assesses how well an arbitrary monotonic function could describe the relationship between two variables, without making any assumptions about the frequency distribution of the variables. Unlike the Pearson product-moment correlation coefficient, it does not require the assumption that the relationship between the variables is linear, nor does it require the variables to be measured on interval scales; it can be used for variables measured at the ordinal level.

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 Time 2 the second round of the Delphi (indicating consensus between the measures), as well as relative rankings that were correlated at the beginning and end of the Delphi (i.e., at Time 1 and Time 2). In addition, the health plan services topic was significantly correlated at the plan level across all four sets of correlations. Access to care had significant correlations across time for its relative rankings (but not for its ratings), as well as between the ratings and relative rankings at the end of the Delphi at Time 2. Provider availability, on the other hand, only had relative rankings and ratings that were correlated at the beginning of the Delphi, at Time 1, but not at the end of the Delphi at Time 2 or across time (at Time 1 and Time 2). 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 availability of provider was ranked highly across all three of the organizational levels by the health plan representatives, we kept it as a high-priority topic over provider communication (which had higher correlations among the ratings and rankings but was not ranked highly by the representatives at an absolute level across any of the three organizational levels; it was ranked a 3 only for the group level).

SUMMARY

Feedback from interviews with health plan representatives confirmed that the QI supplemental survey questions would be useful, so we embarked on documenting the level of importance of particular topics for health plans' QI efforts. We chose a methodology that offered an open but structured forum for plan representatives to share their opinions with us and to rank the topic areas. We conducted a Delphi process with 18 representatives from 13 of the 27 health plans that we had interviewed. Through the Delphi process, we sought consensus on which of the CAHPS health plan survey topic areas were the highest priority for developing additional survey items to provide more data for quality improvement actions.

Based on the results from the Delphi process, the three following topic areas were identified as priority components of the CAHPS health plan survey from the perspective of health plans' information needs for their quality improvement activities:

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

These three priority topic areas became the focus for the next step in our process for development of CAHPS supplemental QI survey items. Working with them, we conducted several sessions with technical review panels to obtain guidance from the participating health plans about specific topic areas within these topic areas for which QI supplemental survey items should be developed. The details of this process are laid out in the next chapter.

CHAPTER FOUR

DEVELOPING SPECIFIC ITEMS WITHIN THE PRIORITY TOPIC AREAS

To gain feedback and guidance from health plans regarding survey items within the priority topic areas, we held a meeting with each one of three different technical review panels of health plan representatives who had participated in both rounds of the Delphi process.

THE REVIEW PANEL PROCESS

The first review panel meeting was held on October 14, 2003, using a Web-based meeting format. Representatives from four of the five health plans invited participated. In addition, representatives from TRICARE –the Tri-Service managed care program that provides all health care for DoD beneficiaries within a DoD geographical region, AHRQ, NCQA, and BCBSA participated as observers, as did representatives from other members of the CAHPS consortium (the AIR and Harvard teams). The health plans invited to the first panel meeting had used CAHPS in conjunction with other internal surveys, and they had offered numerous suggestions and input regarding modification of CAHPS to gear it more toward quality improvement actions. They were among the health plans that reported in our interviews that they used the CAHPS health plan survey primarily because of accreditation requirements. The plan representatives favored developing new survey items for common QI efforts, but they did not support use of CAHPS at the provider level or as a visit-based survey.

These health plans had been performing well on CAHPS quality measures, as reflected in the Quality Compass database. These health plans were sophisticated users of consumer-reported data, and they had already compensated for the limitations of CAHPS for QI by designing and relying on other survey efforts to generate data that could be turned into actions that was not available from CAHPS. Most of them administered their own provider-level survey, as well as various other targeted surveys to consumers. They used these surveys to monitor, identify, and intervene in quality improvement activities.

The second review panel meeting was a telephone conference held on November 14, 2003, to which we invited representatives from five other health plans; all those invited participated. These plans also used CAHPS in conjunction with other internal surveys, and they had offered substantial suggestions regarding modification of CAHPS for QI use. They also favored developing items for common QI efforts. In contrast to the first group, these plans favored use of CAHPS at the provider level or as a visit-based survey, and they did not report that accreditation was their sole reason for using CAHPS. They all praised CAHPS for the scientific integrity of the instrument, its specificity, and the topics addressed in the survey.

The second group of plans had been performing in the medium range on CAHPS quality measures, as reflected in Quality Compass data. Although these health plans used other data sources in addition to CAHPS, they had only partially compensated for the limitations in CAHPS by designing and relying on other survey data collection efforts.

The third review panel meeting was a telephone conference held on November 17, 2003, and another four health plans were invited to participate in this session. Two of the plans participated in the session. These health plans used only CAHPS as their survey data source. They had provided little input in the interviews regarding modifications of CAHPS for QI use. These plans varied in their views about developing items for common QI efforts, but they all favored using CAHPS at the

provider level or as a visit-based survey. The plans varied in their performance on CAHPS quality measures, as reflected in Quality Compass data.

Given the three priority topic areas identified in the Delphi process, the RAND QI and survey team members reviewed the content of the topic areas and the breadth of information within each topic and decided to reframe the three topic areas into six topic areas, for several reasons. The access to care topic was broadened to include availability of providers because of the similarity of content in both topic areas. The health plan services topic, on the other hand, was broken up into four more-specific topic areas because of the breadth of information covered by health plan services. Coordination of care was added as a topic, given that it had been a topic for which developing reliable and useful survey items had been a challenge. These review panels were viewed as a prime opportunity to develop and discuss survey issues within a given topic. In the end, the following six topic areas were presented to the first meeting of 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 had 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).

Even with the prioritizing and narrowing of the number of topic areas (from the original nine topic areas) to those listed above, the range of possible topic areas that could be developed into supplemental QI items remained large. Therefore, the tasks of the first review panel were to (1) reduce the number of topic areas to a manageable number of the most important and (2) identify sub-topics and survey items for each priority topic area they had identified. In the discussion of the items, we asked panelists to consider which level (health plan, group, or individual clinician) they preferred to focus on for item development in each topic area.

In preparation for the first review panel, RAND identified topic candidate items in each of these key topic areas from several sources: existing CAHPS health plan surveys, CAHPS test-item sets (i.e. other CAHPS items that were undergoing development and field-testing), and non-CAHPS surveys. For many of the topic areas and candidate items, RAND had also identified some possible specific, focused action-oriented topics from information gathered from the semi-structured health plan interviews or from previous focus groups concerning these survey topic areas. During the first review panel discussion, these candidate items, by topic and possible focus action-oriented topics, were used as starting point for the panel's work on identifying new, more action-oriented CAHPS items.

The panel was able to identify items that would obtain more focused data on the specific actions needed to be taken for improvement for each of the identified priority topic areas. Through their perspectives and expertise, the health plan representatives on the first panel outlined clear directions and definitive decisions concerning the choices and design of supplemental QI items. Through a similar discussion process with the other two groups of health plans, we were able to assess the extent of consensus on these choices with respect to use of CAHPS for three very different types of health plans.

RESULTS OF THE THREE REVIEW PANELS

The First Technical Review Panel

As a first exercise, the health plans at the first review panel ranked the topic areas relative to each other and offered their reasons for the ranking. Next, they ranked the sub-topics within each of the priority topic areas and defined the recommended organizational level or levels at which each topic should be examined. The health plans chose the following four topic areas as important to further develop for quality improvement: coordination of care, access to care, information and materials, and customer service. The rankings of topic areas and sub-topics within each of the topic areas that resulted from the first technical review panel are presented in Table 4.1.

Coordination of care was chosen as the top topic area by health plans because the representatives of the plans thought it was important to gain the members' perspective on these issues. Coordination-of-care data from consumers is important to health plans, and CAHPS is the only source of the consumer perspective on this issue. Health plans indicated they had other data sources on coordination of care, but none that could provide them with the consumer perspective, particularly in the areas of communication across providers. Health plans were also interested in having additional CAHPS questions in the area of coordination of care.

Access to care was chosen as the second priority topic area. The health plans expressed that access data from consumers is important, and that they need the CAHPS data because of the unique perspective it provides. The plans also had processes in place that integrated the CAHPS access data into other data that they collected. They strongly expressed their desire to keep and not change the CAHPS questions as originally written, because those questions track performance over time and provide benchmarking functions, and they rely on the ways in which the CAHPS measures are integrated into their other mechanisms for quality improvement. However, they were interested in adding questions to focus further in this area. They also felt that all items that are developed and tested on certain important topic areas should be made available for any health plan to include on their CAHPS survey permanently or as a short-term supplemental item. They thought that such inclusion would be especially useful for plans that do not have the data sophistication or monitoring systems that they, as more data-savvy health plans, had in place.

Information and materials was ranked as the third priority topic area. The health plans indicated that they needed the members' perspective on these issues. The CAHPS information for this topic would be more useful if it was broken down into more-specific detail. For example, including specific reference to what source a member used to obtain information-i.e., written materials, Internet.

Customer service was ranked fourth. The health plans indicated that they had other data on customer service, but that the consumer's perspective was important to include. Data on customer service were considered more important than data on claims and paperwork or complaints and appeals.

Table 4.1
Ranking of Topic Areas and Sub-topics by the First Technical Review Panel

Topic Area and Sub-topic	Topic Area Ranking	Sub-topic Ranking	Relevant Level of Service		
			Health Plan	Group	Clinician
1. Coordination of care	1				
a. Communication across providers		1	X	X	X
b. Specific medical conditions		2	X		
c. Getting equipment/ treatment/medication		3	X		
d. Pharmacy services		Deleted	X		
e. Illness prevention		Deleted	X		
2. Access to care	2				
a. Appointments for routine care		1	Added	X	X
b. Appointments with specialists		2	Added	X	X
c. After-hours calls		3	X		X
d. The time it takes to get into a routine office visit, preventive health care visit, and urgent visit		4	X	X	
e. Getting providers		Deleted			
f. Approvals		Deleted			
g. Reasonable waiting times (emergent or urgent)		Deleted			
3. Information and materials	3				
a. Effectiveness of information provided to consumer		1	X		
b. Usefulness of online information and services		2	X		X
c. Use of online information to make decisions about a plan or primary care physician		3	X		
d. Whether consumer got his/her membership ID on time		Deleted			
e. Acceptability of out-of-pocket costs to consumer		Deleted			
4. Customer service	4				
a. Resolution of issues by telephone customer service		1	X		
c. Representative's knowledge and effectiveness		2	X		
b. Representative's politeness and giving of his/her name		3	X		
c. Transportation		Deleted			
d. Pharmacy phone line		Deleted			

NOTE: The topic areas that the health plans did not deem important are ranked as "Deleted."

The Second Technical Review Panel

In the second and third review panel, we presented the priority rankings and the results from the Delphi process and the first review panel meeting. Then we asked the participants for their feedback and opinions on the importance of the topic areas, the importance of sub-topics within each topic area, and specific items that would obtain more focused data on the specific actions needed to be taken for improvement. We convened these two panels because they represent the majority of health plans in terms of data sophistication and CAHPS usage.

The health plans in the second review panel agreed with the importance of the four topic areas and echoed the reasons of the first review panel, but reordered their overall importance, as follows:

1. Access to care
2. Coordination of care
3. Customer service
4. Information and materials.

The reordering occurred because this panel considered access to care an area over which they have very little control and in which they have been working to improve quality. Additionally, they stated that, without access, there is no coordination of care, so they saw it as being important. With respect to customer service being more important than information and materials, the health plans stated that customer service is an important area and one they are constantly trying to improve. The results of the second technical review panel deliberations are reported in Table 4.2.

Table 4.2
Rankings of Topic Areas and Sub-topics by the Second Technical Review Panel

Topic Area and Sub-topic	Topic Area Ranking	Sub-topic Ranking	Relevant Level of Service		
			Health Plan	Group	Clinician
1. Coordination of care	2				
a. Communication across providers		1	X	X	X
b. Specific medical conditions		2	X		
c. Getting equipment/treatment/medication		3	X		
d. Pharmacy services		Deleted	X		
e. Illness prevention		Deleted	X		
2. Access to care	1				
a. Appointments for routine care		1	Added	X	X
b. Appointments with specialists		2	Added	X	X
c. After-hours calls		3	X		X
d. The time it takes to get into a routine office visit, preventive health care visit, and urgent visit		4	X	X	
e. Getting providers		Deleted			
f. Approvals		Deleted			
g. Reasonable waiting times (emergent or urgent)		Deleted			
3. Information and materials	4				
a. Effectiveness of information provided to consumer		1	X		
b. Usefulness of online information and services		3	X		X
c. Use of online information to make decisions about a plan or primary care physician		4	X		
d. Whether consumer got his/her membership ID on time		Deleted			
e. Acceptability of out-of-pocket costs to consumer		2			
4. Customer service	3				
a. Resolution of issues by telephone customer service		1 or 3	X		
c. Representative's knowledge and effectiveness		2	X		
b. Representative's politeness and giving of his/her name		3 or 1	X		
c. Transportation		Deleted			
d. Pharmacy phone line		Deleted			

NOTE: The topic areas that the health plans did not deem important are ranked as "Deleted." The rankings of topic areas and sub-topics that differed from those of the first review panel are boldfaced and shaded in grey.

The health plans at the second review panel agreed with the first review panel on the importance of the topic areas for specific analyses and items that would obtain more focused data on the specific actions needed to be taken for improvement in two of the topic areas—access to care and coordination of care. For the information and materials and customer service topic areas, however,

they had different opinions about the order of importance of the specific topic areas. For the information and materials topic, they added back in a sub-topic that had been deleted by the first panel—acceptability of out-of-pocket costs to consumers. The differences in rankings of topic areas and sub-topics relative to those of the first group are boldfaced in Table 4.2.

The Third Technical Review Panel

The health plans in the third panel were plans that used only the CAHPS health plan survey (not supplementing it with other surveys). They were presented with the results from the Delphi process and the first panel. At this panel’s review meeting, the plan representatives agreed with the importance of the four topic areas, but also reordered the topic areas’ overall importance in the same way that the second review panel did.

The third panel agreed with the first review panel regarding the order of importance of the sub-topics within three of the four topic areas. Within coordination of care, however, they reordered the first and third important sub-topics to those of the first panel, preferring “getting equipment/treatment/ medication” as the most important sub-topic and “communication across providers” as the third ranked (after specific medical conditions, which was ranked second). These health plan representatives thought that getting equipment, treatment, or medication affected the patient’s quality of life more than did communication across providers. In addition, they believed that they could access other data sources for information about communication across providers.

Overall, the three review panels achieved consensus on key aspects and identified four priority topic areas: access to care, coordination of care, information and materials, 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 tracking health plan performance and quality improvement work. In addition, the three panels provided valuable feedback and guidance about the actual wording and content of specific survey items. The next step for RAND was to develop a final list of survey items to be field-tested. The list was developed by compiling the detailed comments, including specific wording, for all of the items selected by the health plans through the consensus process.

CHAPTER FIVE

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 3.0 health plan survey. Currently, the final set of supplemental QI items is being updated for use with the CAHPS 4.0 health plan survey. Through the research steps described in the previous chapters, we were able to identify the topic areas and specific sub-topics for which health plans wanted CAHPS questions that were more oriented to quality improvement actions. Incorporating the detailed comments and guidance from the three review panels, we developed a set of 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 was 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 the supplemental QI items. 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 defined. In some cases, items were allowed to differ from the CAHPS format when that difference was needed to preserve their ability to collect the action-oriented data desired by the health plans. For example, the CAHPS format is generally to have a patient rate an aspect of their care, i.e., rate how often it was easy to understand the information on how a health plan works. For the QI items, a question was designed, for example, to ask what kind of information was *not* easy to understand, with a list of responses to choose from, instead of asking specifically about each individual source of information.

Because the health plans from which we obtained feedback varied widely in their information needs and ratings of importance of items, we anticipated that the needs and priorities of the larger population of health plans also would vary. Therefore, we took an inclusive approach to identification of survey items, developing a menu of supplemental items from which health plans could choose for meeting their specific needs. See Appendix C for a list of the QI items and Appendix D for the actual items in their survey format.

FIELD-TESTING OF THE SUPPLEMENTAL QI ITEMS

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. In this chapter, we present the results of the field test of the QI items, and we offer guidance to health plans on how best to use the items as part of their CAHPS survey data-collection processes.

A health plan can choose a set of supplemental QI items in a given topic for which it wants to identify additional information on needed QI actions. For example, asking, “What was the main reason you had a problem seeing a specialist?” provides a plan with good additional information to the CAHPS question, “In the last 12 months, how much of a problem, if any, was it to see a specialist that you needed to see?”

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.

Overall, this field-test information allowed us to assess both the psychometric performance of the supplemental QI items and their relationship to the CAHPS composites for which they are intended to provide more focused action-oriented QI information. Our analyses of the survey data for the supplemental QI items addressed the following four questions:

1. How well do the supplemental QI items perform psychometrically?
2. How well do the supplemental QI items discriminate performance across health plans?
3. How unique are the QI items to the CAHPS composites?
4. Do each of the QI items within a topic add unique QI information or are there redundancies among the supplemental items? Does the plan have a choice between the supplemental QI items within a topic?

FIELD-TEST RESULTS

Overall, we found that the proposed items had good psychometric properties. The exceptions were a few items that were applicable to very few people or that had more responses in the “other” response category than would be desirable. The initial results also suggest that the QI items discriminate well among plan products, although additional field testing with a larger number of plans and plan products is needed to verify this conclusion.

Most of the individual QI items correlate moderately with their intended composites, and the items correlate moderately with one another. Specifically, the QI “information and materials” items predict their related composite very well, and only a subset of the items is necessary to predict the related composite. The QI “getting needed care” items predict their related composite moderately well, and, again, only a subset of items is needed to predict the composite. Finally, the QI items for the “getting care quickly” topic accounted for only a small portion of the variation in responses for that composite, indicating that the QI items contain unique information that is not captured by the composite and as a result if improvements are made to the specific actions alluded to by the QI items there may not be a subsequent impact on the performance of the associated composite score.

We present here the findings of our analysis of the QI item field-test data for each of the questions defined for the analysis. For each question, we identify the objective, methods, and results of the analysis, and we provide user guidelines based on those results.

1. How well do the supplemental QI items perform psychometrically?

Objective: To determine whether the responses to the CAHPS composites and supplemental QI items are well distributed by examining ceiling effects, 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; floor effects, which means that the data is consistently near the lowest value on the scale (i.e., at the “floor”) and therefore the responses for the item have little to no variation; skewness of distributions; use of “other” response options; and rates of inappropriate missing data. [Note that there is systematic, or appropriate, missing data because not all of the health plans in the field test included all of the items.]

Analysis: Tabulations, frequencies, and univariate statistics were run on all items and composites for all plans.

Results: As shown in Table 5.1, the proposed supplemental QI items and composites were found to have good psychometric properties. The exceptions were a few items that were applicable to very few people (i.e., those needing after-hours care) or had more responses in the “other” response option than would be desirable (i.e., main reason had a problem getting help from health plan’s customer service). As for 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 because 73 percent of the respondents answered “Usually” the highest value in the response scale, but given the context of the question, such an 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 finding out about after-hours care from its population, this set of items is not very useful because few people seek access to after-hours care.

Three QI supplemental questions that ask about reasons for different problems had large percentages of responses in the “other reason” response option:

- 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?

Additional field-test research is scheduled for these three items to identify which additional response categories are needed. The three items will be revised based on the information being gathered.

Guidelines for users: Based on their psychometric properties, 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. The majority of the supplemental QI items in 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), so these items may not be as useful because few people need after-hours care.

Table 5.1 Psychometric Properties of QI Items, Grouped by Topic Area and Composite

Survey Items	Health Plan	Number Answered	Percentage Answered	Item Type	Range of the Response Scale	Mean	Standard Deviation	Percentage Floor	Percentage Ceiling	Percentage "Other"
<u>Access to needed care</u>										
Q9A	A, B, C	1,400	11	C						20
QA	B, C	732	44	D	0-1	0.12				
QB	B, C	86	5	O	1-3	2.32	0.83	23	56	
QC	B, C	37	2	C						8
Q5A	B, C	1,500	91	D	0-1	0.17				
Q5B	B, C	260	16	O	1-3	2.53	0.72	14	67	
Q5C	B, C	124	7	C						41
<u>Getting care quickly</u>										
Q20A	B, C	1,326	80	O	1-3	2.6	0.60	7	67	
Q20B	C	129	16	O	1-4	2.31	0.79	11	9	
<u>Health plan customer service</u>										
Q40A	A, B, C	2,243	18	C						
Q40B	A, B, C	2,243	18	C						
Q40C	A, B, C	4,175	34	D	0-1	0.56				
Q40D	A, B, C	2,354	19	O	1-4	2.96	0.84	6	28	
Q40E	A, B, C	2,343	19	D	0-1	0.41				
Q41A	C	306	38	O	1-4	3.66	0.60	1	73	
Q41B	C	305	38	O	1-5	3.68	1.25	8	35	
Q42A	A, B, C	1,967	16	C						21
Q42B	A, B, C	1,872	15	O	1-5	2.89	1.33	14	21	
<u>New Topic Area: Coordination of care</u>										
Q34A	B, C	1,461	88	D	0-1	0.53				
Q34B	B, C	760	46	O	1-4	3.06	0.98	11	41	
A34C	B, C	753	46	D	0-1	0.64				
Q34D	B, C	487	29	C						
Q34E	B, C	478	29	O	1-5	4.19	0.91	4	41	
<u>Modified Existing CAHPS Items That Were Field-tested</u>										
Q20	A, B, C	9,570	78	O	1-7	3.57	1.62	11	6	
Q40	A, B, C	4,869	40	O	1-3	2.42	0.68	11	54	

KEY: A = Plan 1, B = Plan 2, C = Plan 3. The shading indicates that a particular metric is not applicable to the item within the row.

An item that was fielded by all three plans, indicated by "A, B, C" in column 2, has a sample of 12,195 and 0 systematically missing.

An item that is fielded by two of the plans, indicated by "B, C" in column 2, has a sample of 1,654 and 10,541 systematically missing.

An item that is fielded by one of the plans, indicated by "C" in column 2, has a sample of 814 and 11,381 systematically missing.

Items are "systematically missing" if a plan did not choose to field the item and is therefore missing for all observations in a plan.

Item Type refers to: C = categorical, which means according to or using categories; D = dichotomous, which means having only two values; and O = ordinal, which means having variable classifications that have ordered values.

2. How well do the supplemental QI items discriminate performance across health plans?

Objective: To examine how the supplemental QI items and composites vary across plans and plan products.

Analysis: A staged regression for each composite and QI item was run with dummy variables for the plans (omitting one plan) and dummies for products provided within a plan (omitting a product in each plan). A partial *F*-test was performed that compared the two *R*-squares for each outcome to determine whether there was variation by plan or by product. The reliability of the items at the plan-product level was generated from the regression models, using the Spearman-Brown prophecy formula¹. With data for only three plans, this is a very rough estimation of the plan-level reliability and is not definitive. There is, however, definitive data on the reliability of items at the plan-product level--e.g., comparing a preferred provider organization (PPO) for Plan 1, a PPO for Plan 2, and an HMO for Plan 2.

Results: Table 5.2 reports the results of regressions, each with a QI item as the dependent variable and the health plan and plan products as the independent variables. The responses to the supplemental QI items varied by plan product. Many of the items were compared across only a few products, as shown by the small degrees of freedom reported in the table.

For all but four items, the reliability measures were high and in the reliable range (the values for the reliability measure can range between 0 and 1), showing that the items discriminated performance across health plan products. The four supplemental QI items that did not discriminate across health plan products were QB in access to care, Q41A and Q41B in health plan customer service, and Q34E in coordination of care.

We also conducted these analyses using only the health plan as the independent variable; however, with only three plans fielding the items, there were not enough plans to generate definitive results at the plan level.

Guidelines for users: The supplemental QI items, except questions QB, Q41A, Q41B, and Q34E, are able to discriminate variation in performance across product and have high reliability. The exceptions read as follows:

- 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?

¹ The Spearman-Brown prophecy formula or Spearman-Brown prediction formula relates reliability to test length. The formula is derived from criterion validity calculations; but it can also be used to estimate what the reliability would be if the number of items in a test would be different, or if the test length were different.

Table 5.2
Regression Results with QI Items as the Dependent Variables and
Plans with Products as Independent Variables

Topic Area and Item	Item Type	Plans with Products			
		Degrees of Freedom	<i>F</i> stat	<i>P</i> value	Reliability
<u>Access to needed care</u>					
Q9A	C				
QA	D	3	4.29	0.00	0.76
QB	O	3	0.35	0.78	0.00
QC	C				
Q5A	D	3	3.32	0.01	0.69
Q5B	O	3	2.02	0.11	0.50
Q5C	C				
<u>Getting care quickly</u>					
Q20A	O	3	49.13	0.00	0.97
Q20B	O	1	3.80	0.05	0.73
<u>Health plan customer service</u>					
Q40A	C				
Q40B	C				
Q40C	D	27	6.31	0.00	0.84
Q40D	O	27	2.05	0.00	0.51
Q40E	D	27	3.59	0.00	0.72
Q41A	O	1	0.24	0.62	0.00
Q41B	O	1	0.00	0.99	0.00
Q42A	C				
Q42B	O				
New Topic Area:					
<u>Coordination of care</u>					
Q34A	D	3	2.64	0.04	0.62
Q34B	O	3	15.35	0.00	0.93
A34C	D	3	6.73	0.00	0.85
Q34D	C				
Q34E	O	3	0.83	0.47	0.00

NOTE: Results are not applicable for categorical variables; such variables are shaded to indicate that they are not applicable.

Item Type refers to C = categorical, which means according to or using categories; D =dichotomous, which means having only two values; and O = ordinal, which means having variable classifications that have ordered values.

3. How unique or similar are the QI items to the CAHPS composites?

Objective: To examine correlations among CAHPS composites and supplemental QI items.

Analysis: A correlation matrix was run for the CAHPS composites and supplemental QI items. Four possible results could be found, each of which has a different implication for use of the QI items by health plans in their CAHPS surveys:

1. A QI item could be highly correlated with its relevant composite. If so, a health plan would not have to include such a QI item in the survey because its strong correlation means that the composite score already highly reflects this item's content; therefore, additional direct measurement of the topic via the QI item would not be necessary.
2. A QI item could be moderately correlated with its relevant composite. This item would be most useful for a health plan to include in the survey, because it would be necessary to obtain data on the item to determine whether performance on the item needs improving. If so, such improvement also would help increase scores for the relevant CAHPS composite.
3. A QI item could have a very low correlation with its relevant composite. This would not be a useful item for health plans to use for improving scores on that composite, because it has little effect on the composite. However, the item could be useful to a health plan that is specifically interested in making improvements in the area measured by that item.
4. A QI item is cross-correlated with other composites. If an item is correlated with other composites, as well as the composite to which it is relevant, then a health plan might use it to improve the performance of multiple composites.

Results: Table 5.3 reports the correlations of the CAHPS composites and QI items. The bolded correlations in the table indicate the correlations between the QI item and its *a priori* hypothesized composite. The majority of the candidate items had moderate correlations with their intended composites, and thus should provide useful information for quality improvement efforts. The 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:

- 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 a 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). These 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?

We also found that the "customer service" items are correlated with "after-hours care" items.

Table 5.3 Correlations of CAHPS Composites and Quality Improvement Items

Variable	Label	Item Type	Correlation Coefficients		
			Access: Needed Care	Access: Care Quickly	Health Plan: Customer Service
<u>Access to needed care</u>					
Q9A	Main reason not easy to get specialist appt	C			
QA	Need visit dr office for after-hrs care	D	0.11*	0.04	0.03
QB	Problem getting needed after-hrs care	O	0.60*	0.35*	0.39*
QC	Main reason problem getting after-hrs care	C			
Q5A	Call personal dr after hours	D	0.04	-0.01	0.01
Q5B	Problem talk to personal dr after hrs	O	0.30*	0.43*	0.18
Q5C	Main reason problem talk to personal dr after hr	C			
<u>Getting care quickly</u>					
Q20A	Problem wait as long as did to see dr	O	0.37*	0.43 *	0.20*
Q20B	Problem set up appt due to dr schedule	O	0.07	0.33 *	0.01
<u>Health plan customer service</u>					
Q40A	Type of info had problem understand or use	C			
Q40B	Where got info had problem understand or using	C			
Q40C	Look for info on plan Internet site	D	0.07*	0.06*	0.03
Q40D	How useful plan Internet site info	O	0.27*	0.22*	0.58*
Q40E	Use plan Internet site info to choose dr	D	0.08*	0.06*	-0.08*
Q41A	Customer svc staff courteous respectful	O	0.35*	0.27*	0.43*
Q41B	How certain customer srvc would get you help	O	0.14	0.14	0.44*
Q42A	Main reason problem get help from customer srvc	C			
Q42B	Number calls needed to get customer srvc help	O	0.11*	0.03	0.38*
<u>Topic Area without a CAHPS composite: Coordination of care</u>					
Q34A	Get care from more than one kind of dr	D	0.03	0.04	0.05
Q34B	Dr informed about care from other drs	O	0.27*	0.40*	0.11
A34C	Office coordinates care among diff drs	D	-0.02	-0.10	-0.01
Q34D	Who coordinated care in last 12 months	C			
Q34E	How satisfied with coordinated care help	O	0.30*	0.26	0.16

* Item is statistically significant at the $p < 0.05$ level.

NOTES: The **bolded** correlations indicate that the QI item is associated with its *a priori* hypothesized composite. Both the “coordination of care” items and the entire “coordination of care” composite are supplemental and therefore do not have an *a priori* hypothesized composite. Correlations reflect coding in which “high” is always good for both candidate QI items and composites. These results are not applicable for categorical variables and are therefore shaded when not applicable. Also, abbreviated item content is used for the label.

Item Type refers to C = categorical, which means according to or using categories; D = dichotomous, which means having only two values; and O = ordinal, which means having variable classifications that have ordered values.

Guidelines for users: 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 and therefore would not be helpful for health plans to use 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 were 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.

4. Does each of the QI items within a topic add unique QI information, or are there redundancies among the supplemental items? Does the plan have a choice between QI items within a topic?

Objective: To further test relationships among the supplemental QI items using regression analyses. If any of the supplemental QI items are highly correlated with each other, they may be redundant and a choice can be made between either of the items, according to the plan’s preference. If the QI items are not correlated with each other, then each item is providing unique information about the consumer’s experience, and plans can gain increased information by including more of the items in the survey.

Analysis: Correlations were estimated among the supplemental QI items, and multivariate regressions were run for each composite. “Coordination of care” supplemental items were not included in this analysis because these items do not have a current, corresponding CAHPS composite, so there was no previous CAHPS composite to correlate them with. The unit of analysis for the regressions was individual-level responses to the survey (one data record for each respondent). In each regression, the dependent variable was the individual-level composite score, and the independent variables were response values for the supplemental QI items related to the composite. In addition, dummy variables were included that coded for missing data on each item. For supplemental QI items that a respondent did not answer, we assigned the mean value of the variable and used these imputed values for the analysis. In addition, dummy variables were created for items that had systematic missing data because not all the health plans fielded the items. These dummies for missing data were included in the regressions so that all of the data across the three plans could be used in all the regressions, maximizing our use of the survey information.

Results: The supplemental QI items for previously defined CAHPS composites are correlated with each other, but the correlations are moderate enough that there are no redundancies among the items. The Table E.1 in Appendix E reports the correlations among the QI items. There are no correlations above 0.80, which would indicate that items are redundant in content. Because all the items are fairly independent, each of them is adding some unique information.

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. Table 5.4 provides the results from the regressions, with the composites as the dependent variable

and the *a priori* (intended) corresponding set of QI items as independent variables. Bolded items are statistically significant. Table 5.5 reports regression results, with the composites as dependent variables and the *a priori* corresponding set of QI items as independent variables, along with the results from the stepwise regression, which includes the QI items in blocks. In Table 5.5, we boldface the theorized items that contribute to the composite, as well as the model results that are significant and contribute to the prediction of the composite.

These regression analyses 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:

- As a single item, Q9A, “What was the main reason you had a problem seeing a specialist?” was moderately predictive (R^2 of 0.22) of the “getting needed care” composite, and was the majority of the predictive power of the *a priori* set (which combined had the predictive power of an R^2 of 0.23)
- The *a priori* set of two items of the “getting access to care quickly” composite was not really predictive (R^2 of 0.02). These were Q20A, “In the last 12 months, how much of a problem, if any, was it to see a provider” and Q20B, “In the last 12 months, how often did you have a problem getting an appointment because the provider you wanted to see worked limited hours or had few available appointments?”
- Three items were highly predictive (R^2 of 0.60) of the “health plan customer service” composite. These were the combined *a priori* set: 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: Each of the supplemental QI items offers unique information to the composites to which the item is related. Therefore, all of those items with moderate correlations to the composites could generate usable information for QI if they were included on the survey.

As for predicting performance on the composites, a smaller number of key items with the strongest influence on each composite 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.

More specifically, a health plan could minimally include the following supplemental QI item or items for the given composite:

“Getting needed care”:

- Q9A, What was the main reason you had a problem seeing a specialist?

“Getting care quickly”:

- Q20A, In the last 12 months, how much of a problem, if any, was it to wait as long as you did to see a provider?

“Health plan customer service”:

- Q42B, How many calls did it take for you to get the help you needed from your health plan’s customer service? or
- Either Q42A, What was the main reason you had a problem getting help from your health plan’s customer service? And/or Q40A, What kind of information did you have a problem understanding or using? (Specifically, the choice—getting care outside of your network).

Table 5.4
Regression Results, with Composites as Dependent Variables and the *A Priori* Corresponding Set of Quality Improvement Items as Independent Variables

<i>A Priori</i> Set of Variables	Total Model <i>R</i> ²	Coefficient	Standard Error
<hr/>			
<i>Access: Getting needed care</i>	0.23		
QA: Need visit dr office for after-hours care		-7.20	14.8
QB: Problem getting needed after-hours care		7.48	6.7
QC: Main reason problem getting after-hours care		0.57	2.0
Q5A: Call personal dr after hours		0.60	12.3
Q5B: Problem talk to personal dr after hours		7.90**	2.4
Q5C: Main reason problem talk to personal dr after hours		0.75	1.6
Q9A_2: Problem see spec: Auth delayed/denied		-5.22*	2.6
Q9A_3: Problem see spec: Couldn't find list of spec		14.25***	4.0
Q9A_4: Problem see spec: Had to choose spec far away		7.98	3.3
Q9A_5: Problem see spec: Not enough spec choose from		5.77	3.3
Q9A_6: Problem see spec: Spec wanted not in HP		2.92	2.5
Q9A_7: Problem see spec: Appt time not convenient		11.33***	2.5
Q9A_8: Problem see spec: Other reason		6.75**	2.5
<hr/>			
<i>Access: Getting care quickly</i>	0.02		
Q20A: Problem wait as long as did to see dr		14.3***	1.2
Q20B: Problem set up appt due to dr schedule		7.87**	2.6
<hr/>			
<i>Health Plan customer service</i>	0.60		
Q40Aa : Problem info: Dr specialist visits		-1.82*	0.9
Q40Ab: Problem info: Pharmacy		-2.2*	0.9
Q40Ac: Problem info: Specialist referral		-1.87*	1.1
Q40Ad: Problem info: After-hrs/ER care		1.89	1.5
Q40Ae: Problem info: Choosing health provider		-3.00*	1.2
Q40Af: Problem info: Getting care outside network		-3.55**	1.1
Q40Ag: Problem info: Something else		-0.58	1.1
Q40Ba : Problem info source: Employer		2.39	1.2
Q40Bb: Problem info source: Plan Internet site		-0.38	1.1
Q40Bc: Problem info source: Plan call center		11.43***	0.9
Q40Bd: Problem info source: Plan written info		0.93	1.0
Q40Be: Problem info source: Plan sales rep		-8.02***	2.2
Q40Bf: Problem info source: Drs office		2.82*	1.2
Q40Bg: Problem info source: insurance agent		-0.67	2.7
Q40Bh: Problem info source: Some other source		-0.14	1.4
Q40Bi: Problem info source: Not sure		-0.35	1.6
Q40C: Look for info on plan Internet site		-1.39	5.9
Q40D: How useful plan Internet site info		2.95***	0.5
Q40E: Use plan Internet site info to choose dr		-1.75**	0.8
Q41A: Customer srvc staff courteous/respectful		0.78	2.0
Q41B: How certain customer srvc would get you help		1.22	1.0
Q42B: Number of calls to get customer srvc help		-5.57***	0.3

* Item is statistically significant at the $p < 0.05$ level; ** item is statistically significant at the $p < 0.01$ level; *** item is statistically significant at the $p < 0.001$ level.

NOTES: All models include the missing flags for the independent variables in the model. Also, abbreviated item content is used for the label.

Table 5.5
Regression Results for Composites as Dependent Variables, Using Stepwise Regression to
Include QI Items in Blocks as Independent Variables

Variables	Total Model R ²	Incremental R ²
<i>Access: Getting needed care</i>		
<i>A Priori Set:</i> QA, QB, QC, Q5A, Q5B, Q5C, Q9A	0.23	--
Stepwise:		
Q9A dummies: Main reason problem seeing specialist	0.22	0.22*
Q40Aa–Q40Ag: Kind info problem understand/use	0.25	0.02*
Q42B: Number calls needed to get customer srvc help	0.26	0.01*
Q20A, Q20B: Problem wait see dr, Problem set up appt	0.27	0.00*
Q40C, Q40D, Q40E: Used Internet site, Internet site useful, Used Internet to choose dr	0.28	0.00*
Q40Ba–Q40Bi: Where got info had problem understd/use	0.28	0.00*
Q42A: Main reason problem getting help customer svc	0.28	0.00
<i>Access: Getting care quickly</i>		
<i>A Priori Set:</i> Q20A, Q20B	0.02	--
Stepwise:		
Q9A dummies: Main reason problem seeing specialist	0.03	0.03*
Q20A, Q20B: Problem wait see dr, Problem set up appt	0.05	0.01*
Q40Aa–Q40Ag: Kind info problem understand/use	0.06	0.01*
Q42B: Number calls needed to get customer srvc help	0.07	0.00*
Q40C, Q40D, A40E: Use Internet site, Internet site useful, Used Internet to choose dr	0.07	0.00*
Q34A, Q34B, Q34C, Q34Da–Q34De: More than one provider, Dr informed about other drs, offc coordinate care among diff drs, Who coordinated care, Satisfied with coordinate care	0.08	0.00*
Q5A, Q5B, Q5C: Need visit dr after hours, Problem getting after-hours care, Main reason problem getting after-hours care	0.08	0.00*
Q40Ba–Q40Bi: Where got info, problem understand/use	0.08	0.00
<i>Health Plan customer service</i>		
Variables:		
<i>A Priori Set:</i> Q40Aa–Q40Ag, Q40Ba–Q40Bi, Q40C, Q40D, Q40E, Q41A, Q41B, Q42B	0.60	--
Stepwise:		
42B: Number calls needed to get customer srvc help	0.36	0.36*
Q40Ba–A40Bi: Where got info had problem understand/use	0.58	0.21*
Q40C, Q40D, Q40E: Use Internet site, Internet site useful, Used Internet to choose dr	0.60	0.02*
Q40Aa–Q40Ag: Kind info problem understand/use	0.60	0.00*
Q42A: Main reason problem getting help customer srvc	0.60	0.00
Q41C: How certain customer srvc would get you help	0.60	0.00

*Model results that are significant and contribute to the prediction of the composite.

NOTE: Theorized items that contribute to the composite are **bolded**. Abbreviated item content is used for the label.

CHAPTER SIX DISCUSSION AND CONCLUSION

As the use of the CAHPS health plan survey increased, health plans began to express concerns that the survey did not provide them with action-oriented information that could help them improve performance in areas important to consumers, as reflected in the survey results. Both the National Committee on Quality Assurance and the Blue Cross Blue Shield Association had received feedback from their constituent health plans regarding these same concerns.

In response, RAND pursued a coordinated strategy to strengthen the quality improvement aspects of CAHPS. As part of this strategy, RAND partnered with NCQA and BCBSA to study the use of CAHPS in health plan quality improvement activities, and to develop a set of CAHPS supplemental QI items that health plans can use to collect information on which actions can be taken for quality improvement. As part of this development process, the RAND team assessed the performance of the developed supplemental QI items and provided guidance to users on use of these items to generate information for their QI processes.

The set of QI items was developed for the CAHPS 3.0 health plan survey, using the participative process with health plans, the BCBSA, and NCQA, as described by this report. These items were field-tested in 2004 in partnership with three health plans. Currently, the item set is being updated for use in the 4.0 health plan survey.

IDENTIFICATION OF IMPORTANT ISSUES FOR HEALTH PLANS

In retrospect, we cannot overstate the importance of taking the steps to learn from the health plans, as key stakeholders of the CAHPS survey, about concerns they may have regarding the survey content and process, in particular its implications for their quality-improvement processes. As shown clearly from our health plan interviews, the plans in general respected the strengths and current uses of CAHPS, but they also had concerns about the usability of CAHPS data for QI efforts to improve performance on the survey. The health plans reported that CAHPS was a scientifically sound tool for comparing themselves to other health plans and for examining trends in their performance over time. However, the plans also reported that CAHPS had limited ability to provide useful information for QI because the CAHPS data are reported at the plan level only, the contents of questions are too general to be oriented to particular actions, and results are not timely enough to be useful for monitoring of QI progress.

Feedback from the health plans pointed toward the need to develop new supplemental items in future CAHPS surveys that the plans could use as QI items that would obtain more focused data on the specific actions needed to be taken for improvement activities. Their interview responses indicated that they did not want to change the existing survey content areas; instead, they preferred to have additional items through which they could obtain more action-oriented data on their performance on those content areas. However, a number of plans voiced concerns about how agreement would be reached on selection of QI issues for item development, and some were worried about adding to respondent burden. Overall, health plans did not want the changes to increase costs or survey length, or to result in weakening the credibility, standardization, and comparability that are strong features of the CAHPS health plan survey data.

It was with these preferences and concerns in mind that the RAND team developed a participative approach through which health plans would “drive” a consensus process to identify which supplemental QI items should be developed for the CAHPS survey. This process included decisions regarding (1) which topic areas were the most important ones to address for the largest

percentage of plans and (2) within each topic area, which sub-topics were the most important for development of QI items. Our goal was to ensure that the QI items added to the CAHPS family of survey items would be of optimal usefulness to health plans, within the constraints of limited resources for both item development and survey administration.

The health plan participants in the three review panel achieved a high level of consensus on both the key topic areas and sub-topics within those topic areas for which QI items should be developed, despite differences in their levels of experience in using survey data from CAHPS and other survey tools. There was clear agreement on the four priority topic areas of access to care, coordination of care, information and materials, and customer service. The panels also identified the same sub-topic within these topic areas, although their opinions differed somewhat regarding the order of priority for those sub-topics. In addition, the three panels provided valuable feedback and guidance about the actual wording and content of specific survey items.

USABILITY OF THE SUPPLEMENTAL QI ITEMS

With this guidance from the participating health plans, RAND could proceed with some assurance that the items we developed and tested would be useful for the target audience. The results of the field test for the QI items showed that, with a few exceptions, the items perform well psychometrically and are correlated to the composites for which there are developed items that are aimed to obtain more focused data on the specific actions needed to be taken for improvement. Therefore, health plans can use them with confidence. The QI items for the “getting care quickly,” “health plan customer service,” and “coordination of care” composites, as well as the majority of the QI items for the “getting needed care” composite, had good psychometric properties. However, few patients responded that they needed after-hours care, so the related QI items may not be useful because of the very small percentage of responses to them.

Importantly, we found generally good correlations between the supplemental QI items and the CAHPS composites to which they are related. Therefore, by focusing interventions to improve performance on these QI items, health plans have the potential to improve their scores on the relevant CAHPS composites. At the same time, none of the QI items was so highly correlated with its relevant composite that a health plan would not want to include the QI item in the survey because the composite score already highly reflects this item’s content. Furthermore, the QI items related to each of the “customer service” and “after-hours care” composites also are correlated to performance on the other composite, so that improving performance on one or more of these items could affect performance on both composites.

CONCLUSION

We anticipate that the supplemental QI items developed through this extensive process will be useful to health plans in gathering action-oriented information to help them 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 health plans, 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.

APPENDIX A

DESIGN AND METHODS FOR THE HEALTH PLAN INTERVIEWS

We needed to use a data-collection methodology that offered the flexibility for respondents from health plans to share with us their unique experiences in working with the CAHPS health plan survey. Therefore, we chose to conduct semi-structured interviews with health plan representatives from a purposive sample of 27 health plans described further below. The protocol used was a component of a more general interview protocol that the CAHPS consortium used for broader market research interviews conducted during 2003.

PROTOCOL CONTENT

We developed and field-tested a standardized interview protocol to guide the health plan interviews. Each interview took approximately one hour. The protocol was designed to allow us to:

- Gain an understanding of the importance of consumer measures for health plans
- Identify how the health plans are using CAHPS in general and for quality improvement (QI) activities specifically
- Obtain feedback on the value and limitations of CAHPS for quality improvement,
- Identify topic areas that are important to health plans that could provide more action-oriented data from CAHPS.

The interview protocol included the following main questions:

1. What is your organization most interested in knowing from consumers or patients about their experiences with their health care and customer service? Why? What drives their experiences as patients or consumers of health care?
2. What surveys do you use to collect consumer experience with care? CAHPS? Others? Can you tell me a little about the survey that you use? Audience? Reference period? Frequency of administration? Unit of analysis?
3. Are CAHPS or your other survey(s) useful for quality improvement activities of health plans or providers?
4. Now, thinking about the content of the surveys, what do you find *most useful* about the content of the surveys you use? *Least useful*?
5. To people important to you and your organization, is {CAHPS/your current survey} credible with respect to: a. Scientific integrity of the survey instrument; b. Topic areas addressed by the survey; c. Specificity of the information generated from the survey? Y/N, Why?
6. Does the {CAHPS/your survey/both} measure processes of care or service issues that you think are important? *IF YES*, which ones? *IF NO*, why not? What could be improved or added?
7. Does the {CAHPS/your survey/both} provide you with action-oriented data? *IF YES*, in what areas? How most commonly used? *IF NO*, what would you like to see produced by the CAHPS development team that would make the data oriented toward taking specific actions aimed at improvement within QI activities? [Probe: specific changes or different measures]

8. What data are most useful to your organization for benchmarking or for making comparisons? [PROBE: Local? Regional? National? Other--e.g., by Specialty? Group size? Urban? Rural?]
9. Are data from the consumer-assessment surveys compared to benchmarks such as the National CAHPS Benchmarking Database (NCBD) or the National Committee for Quality Assurance (NCQA) Quality Compass data?
10. What aspects of the CAHPS health plan survey or methods are useful for QI?
11. How do you use CAHPS for QI? For example: to benchmark, to identify problem areas, to establish improvement goals, to develop specific interventions, to monitor intervention effectiveness.
12. What do you think are the key barriers for using CAHPS data for quality improvement?

SAMPLE OF HEALTH PLANS

Interviews were conducted with personnel in a total of 27 health plans. The sample of health plans we interviewed was identified in collaboration with the Blue Cross Blue Shield Association (BCBSA) and the NCQA. To obtain as broad a range of viewpoints as possible, we selected health plans to interview from three categories—plans that did not field CAHPS, plans that fielded CAHPS but did not publicly report CAHPS results in the NCQA’s Quality Compass product, and plans that did publicly report CAHPS results in Quality Compass. Within the group of health plans that report their CAHPS results to Quality Compass, we selected plans with a range of performance levels on the CAHPS health plan survey. Within each of the three sample groups, we attempted to select plans from across the regions of the United States (North, South, East, West, and Midwest).

For each plan, an interview was conducted with the person responsible for management of the QI function. We expected the QI managers or directors to have both operational knowledge of the CAHPS health plan survey and an understanding of how the CAHPS data are used throughout the organization. The interviews were with the QI managers from health plans in each of the following groups:

- BCBS health plans that do not use the CAHPS health plan survey ($N=4$)
- BCBS health plans that field the CAHPS health plan survey but do not report survey results to the NCQA Quality Compass for public reporting ($N=4$)
- Health plans that report survey results to the NCQA Quality Compass for public reporting, grouped by performance on CAHPS in 2000 and 2002, as identified from survey data submitted to NCQA Quality Compass, as follows (total $N=19$):
 - NCQA-1: Plans that consistently had CAHPS scores higher than average ($N=5$)
 - NCQA-2: Plans that consistently had CAHPS scores lower than average ($N=4$)
 - NCQA-3: Plans with mixed performance on CAHPS (not included in interviews)
 - NCQA-4: Plans that had CAHPS scores that showed statistically significant improvement between 2000 and 2002 ($N=5$)
 - NCQA-5: Plans that had CAHPS scores that showed a statistically significant decline between 2000 and 2002 ($N=5$).

HEALTH PLAN RECRUITMENT AND PARTICIPATION

We collaborated with BCBSA in the recruitment of BCBS health plans to participate in the interviews. To initiate the process, the BCBSA fielded a Web-based survey to its member health plans in which it gathered basic descriptive information on the plans' quality improvement activities and whether and how they were using CAHPS. It also asked them to indicate whether they were willing to be interviewed by RAND to obtain their views on CAHPS.

A total of 24 out of 42 BCBS health plans responded to the BCBSA survey, and 23 of these 24 plans stated their willingness to participate in the RAND interviews. All of the BCBS health plans identified through the BCBS Web-based survey were using CAHPS. Of the 23 health plans, 17 plans did not report their CAHPS results to the NCQA Quality Compass. We identified 9 of the 17 plans for possible interviews, selecting plans located in a variety of geographic regions. From this list, 4 plans were interviewed, 3 plans did not respond to calls, and 2 plans were not contacted because an interview already had been scheduled with another plan in their region.

The BCBSA also provided us with the names of the 8 plans that they knew of that did not use CAHPS at all. From this list, 4 plans participated in the interviews, 2 plans did not respond, and 2 plans chose not to participate.

With the analytic support of NCQA, RAND undertook the selection of health plans that reported CAHPS results publicly to NCQA Quality Compass. NCQA Quality Compass is a database of health plans that allows NCQA to publicly release their results. This subset of health plans was used for this RAND CAHPS study. The first step, which was an analysis of the Quality Compass health plans' CAHPS performance, was guided by a written analysis plan prepared jointly by RAND and NCQA. NCQA provided the RAND team with health plan-level summary CAHPS survey data for the years 2000 and 2002. A table was also provided to match the plans that reported results to Quality Compass for all three years. Data were available for a total of 245 health plans in 2002. For a subset of 154 plans that had three years of data, the data were matched and trended from 2000 to 2002. The CAHPS 2.0 performance of each plan was indexed using the standardized difference of the plan's scores from the means of the group. These standardized scores were based on raw scores pre-calculated by NCQA using its standard methods, as documented in the detailed instructions presented in *HEDIS® 2002 Volume 3: Specification for Survey Measures* (NCQA, 2002).

Working with the results of the analysis, the RAND team prepared a list of the health plans that were candidates for interviews. This list included 5 to 10 plans in each of the four performance categories described above: consistent improvers, improvers, decliners, consistent decliners. The NCQA sent an introductory letter to those plans, informing them about the purpose of the interviews and delineating informed-consent provisions. These 30 health plans were asked to contact NCQA if they did not wish to participate.

The NCQA provided RAND with contact information for 27 health plans that did not decline to participate in the interviews. Of these health plans, 19 plans were interviewed, 7 plans across the four performance levels were not contacted because interviews were scheduled with 5 other plans with the same performance level, and 1 plan refused to participate.

A member of the RAND research team conducted each interview by telephone. In some cases, multiple individuals from the same health plan participated in the interview.¹ A total of 27

¹ In some of the interviews, we interviewed more than one person. In 18 of the interviews, we interviewed only one person, the primary contact person. In the remaining nine interviews, we interviewed multiple people at a time. In five of the interviews, we interviewed two people together; in one interview, there were three people; in two interviews,

individual or group interviews were conducted. We used the informed-consent procedure to confirm the earlier consent given by BCBS health plans and to obtain oral consent from the non-BCBS/NCQA health plans. The interviews were audio taped, and the results of each interview were written up based on notes taken during the interview and information from the tapes.

Overall, 44 health plans were recruited, of which 27 plans were interviewed; five plans did not respond; nine plans were not contacted because the number of health plans needed by region was reached; and three plans refused the interview. The interviewees had an average of 4.7 years of experience in their current position and held the positions and job titles listed in Table A.1. In addition, 40 percent ($N=11$) reported to a Vice President, 37 percent ($N=10$) reported to a Director, 11 percent ($N=3$) reported to a Chief Medical Officer, two reported to other positions, and information was missing for one respondent.

Table A.1
Distribution of People Interviewed, by Type of Position Held

Position Title	Number Interviewed (N)	Percentage (%) of $N=26$
Director–QI/Accreditation	3	11
Director–Research/Consulting	3	15
Director–Medical-related depts.	2	7
Associate/Assistant Director	2	7
Project leader/Manager of QI	4	15
Project leader/Manager of Market Research	4	15
Project leader/Manager of other dept	2	7
Analyst/Consultant	6	22
Missing	1	<1

As to the health plans interviewed, they were primarily network models or a combination or mix of models. Tables A.2 and A.3 provide self-reported health plan information about total enrollment (including commercial, Medicare, Medicaid, etc.), model type, and region of the United States.

there were four people; and in one interview, there were five people. In each interview, however, the main contact person was considered the “primary interviewee.”

Table A.2 Total Enrollment Size of the Health Plans Included in Interviews

Group	Health Plan Enrollments		
	Mean	Minimum	Maximum
NCQA-1 (N=5)	373,250	30,000	1,300,000
NCQA-2 (N=4)	957,600	63,000	2,200,000
NCQA-4 (N=5)	936,459	200,000	2,574,836
NCQA-5 (N=5)	3,485,400	217,000	13,400,000
Field CAHPS, No public reports (N=4)	1,400,000	300,000	3,500,000
Don't field CAHPS (N=4)	1,372,500	600,000	3,000,000

SOURCE: Self-reported data from the health plans.

NOTES: Health plans that report survey results to the NCQA Quality Compass for public reporting were grouped according to performance on CAHPS in 2000 and 2002, as identified from survey data submitted to NCQA Quality Compass, as follows (total N=19):

NCQA-1: Plans that consistently had CAHPS scores higher than average (N=5)

NCQA-2: Plans that consistently had CAHPS scores lower than average (N=4)

NCQA-3: Plans with mixed performance on CAHPS (**Not included in interviews**)

NCQA-4: Plans that showed a statistically significant improvement in CAHPS scores from 2000 to 2002 (N=5)

NCQA-5: Plans that showed a statistically significant decline in CAHPS scores from 2000 to 2002 (N=5).

Table A.3 Distribution of Model Types for the Health Plans Included in Interviews

Group	Health Plan Model Type					Total
	IPA	Mixed	Network	Staff	Other	
NCQA-1	--	1	2	1	1	5
NCQA-2	--	2	2	--	--	4
NCQA-4	1	2	2	--	--	5
NCQA-5	2	2	--	--	1	5
Field CAHPS, no public reports (N=4)	1	2	1	--	--	4
Don't field CAHPS (N=4)	1	2	1	--	--	4
Total	5	11	8	1	2	27

SOURCE: Self-reported data from the health plans. Those listed as "Other" did not specify their health plan model.

NOTES: IPA = independent physician association, is when the HMO is organized so that the participating physicians maintain their own separate offices. Such physicians usually treat both private patients and HMO members. Network = network model, which is when a health plan contracts with multiple physician groups or other providers to deliver health care to members (known as approved providers). Mixed = mixed model, which is a combination of an IPA and a network model. Staff = staff model, which is a type of HMO, similar to a group model, in which physicians are salaried employees who provide their services exclusively to HMO enrollees.

Health plans that report survey results to the NCQA Quality Compass for public reporting, were grouped based on performance on CAHPS in 2000 and 2002, as identified from survey data submitted to NCQA Quality Compass, as follows (total N=19):

NCQA-1: Plans that consistently had CAHPS scores higher than average (N=5)

NCQA-2: Plans that consistently had CAHPS scores lower than average (N=4)

NCQA-3: Plans with mixed performance on CAHPS (**Not included in interviews**)

NCQA-4: Plans that showed statistically significant improvement in CAHPS scores from 2000 to 2002 (N=5)

NCQA-5: Plans that showed statistically significant decline in CAHPS scores from 2000 to 2002 (N=5).

ANALYSIS OF THE INTERVIEW DATA

The interviewers drafted notes from the tape recordings of the interviews, and these notes were given to another team member. Working with the notes and the tape, the other team member

drafted a document of the content of the interview. The interviewer then reviewed the draft interview for accuracy and finalized the interview document. These final interview documents were used to identify patterns and to create a detailed spreadsheet of quantifiable concepts for each interview. The spreadsheet included health plan characteristics (i.e., model type, title of person interviewed) and health plan opinions (i.e., main use of CAHPS). The spreadsheet allowed us to do frequency counts and calculations for quantifiable data and aided in sorting and grouping interviews for qualitative analysis. The counts were also disaggregated according to the health plan sample-selection criteria (i.e., do not field CAHPS, do not report CAHPS publicly, and field and publicly report CAHPS, grouped by the four performance categories) to identify whether answers clustered according to these selection characteristics.

APPENDIX B

RELIABILITY OF RATINGS OF IMPORTANCE OF QI ITEMS

We invited 34 individuals from 25 of the 27 interviewed health plans to participate in the Delphi method (and perform the topic rankings and ratings). The two non-included health plans had never fielded CAHPS. Of the 34 individuals, 18 from 13 health plans completed the rankings in both rounds—a 72-percent response rate for health plans, and a 53-percent response rate for individuals.

The 18 health plan representatives provided 0–10 ratings of the importance of nine topic areas—(1) access to care, (2) provider availability, (3) complaints and appeals, (4) provider communication, (5) coordination of care, (6) health plan services, (7) authorization of care, (8) ancillary clinical services, and (9) preventive care—at each of three levels (plan, group, individual provider) at two time points, as part of a Delphi process.

We estimated the reliability of the importance ratings at both time points. We estimated reliability by using a one-way model, partitioning topic areas according to the between-variance factor (9 topic areas \times 3 levels = 27) from the within-variance factor. The estimated reliability of ratings was 0.79 (intraclass correlation = 0.18) at Time 1 and 0.92 (intraclass correlation = 0.42) at Time 2 (refer to Tables B.1 and B.2).

We also estimated correlations between ranks and ratings of the importance of each of the nine topic areas within each level of health care. Spearman rank-order correlations for the plan level at Time 1 and Time 2 are provided in Table B.3 and Table B.4, respectively. We expected these correlations to be negative, because a lower number represents a higher rank, but to have a lower-level rating of importance.

At Time 1, the correlations ranged from -0.19 to -0.82 for health plan topic areas, from -0.15 to -0.62 for medical group topic areas, and from 0.03 to -0.66 for individual clinician topic areas. Six of the nine correlations were statistically significant at the plan level, but only one correlation was significant at the group level and only one, at the individual clinician level. These six health plan–level topic areas were provider availability (-0.64), provider communication (-0.72), health plan services (-0.74), authorization of care (-0.82), ancillary clinical services (-0.49), and preventive care (-0.62).

At Time 2, the correlations ranged from 0.00 to -0.63 for health plan topic areas, 0.07 to -0.57 for group topic areas, and 0.11 to -0.49 for individual clinician topic areas. Only three of the plan topic areas, two of the group topic areas, and none of the individual clinician topic areas were statistically significantly correlated at Time 2. These three health plan–level topic areas were: access to care (-0.52), complaints and appeals (-0.63), and health plan services (-0.63).

Correlations at Time 1 and Time 2 for the group and individual clinician levels are not reported here. However overall, the ratings and ranks were more consistent for the plan level than for the group and individual clinician levels at both Time 1 and Time 2. The consistency between rankings and ratings at all three levels also decreased over time, indicating that the ratings and rankings were picking up different and unique information about the importance of the topic areas.

Next, correlations were computed between the ranks of topic areas within level across the two time points. Table B.5 provides the correlations for the plan from Time 1 to Time 2. Spearman correlations between ranks ranged from 0.33 (availability) to 0.80 (access) for

plan, 0.33 (communication) to 0.89 (complaints) for group, and 0.20 (ancillary) to 0.83 (complaints) for individual clinician. Seven of the nine correlations were statistically significant at the plan level from Time 1 to Time 2: access to care (0.76), complaints and appeals (0.83), provider communication (0.66), coordination of care (0.75), health plan services (0.63), authorization of care (0.52), and preventive care (0.70). This significance indicates that the rankings of these seven topic areas were stable among the health plan representatives over time.

Correlations were also computed between the ratings of topic areas within level across the two time points. Table B.6 provides the correlations for the plan level of health care. Spearman correlations between ratings ranged from 0.28 (availability) to 0.84 (communication) for plan, -.13 (complaints and appeals) to 0.87 (preventive care) for group, and 0.13 (communication) to 0.66 (provider availability) for individual clinician. Five of the nine correlations were statistically significant at the plan level from Time 1 to Time 2: provider communication, health plan services, authorization of care, ancillary clinical services, and preventive care. This significance indicates that the health plan representatives shifted the importance level of topic areas from Time 1 to Time 2.

Moreover, the rankings were consistent from Time 1 to Time 2 at the health plan level for all of the topic areas except availability of providers, and the majority of the ratings of the topic areas were consistent from Time 1 to Time 2, except for access, availability of providers, complaints and appeals, and coordination of care.

In particular, two of the topic areas that were ranked of highest importance by the Delphi--access and health plan services--had ratings and relative rankings that were significantly correlated at the health plan level at Time 2 (indicating consensus between the measures), as well as relative rankings that were correlated 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. Access to care 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. Provider availability, on the other hand, had relative rankings and ratings that were correlated only at Time 1, not in Time 2 or across time. These correlations validated the overall rankings identified by the Delphi participants for the top two topic areas--access to care and health plan services. And, importantly, the rankings of the topic areas that were ranked of highest importance by the Delphi process by the health plan representatives--access, coordination of care, and health plan services--were significantly correlated from Time 1 to Time 2 at the health plan level.

Table B.1
Partitioning Between-Topic-Area Variance and Within-Topic-Area Variance at Time 1

Source	Degrees of Freedom	Mean Square	Label for Mean Square
Between	26	22.58	BMS
Within	448	4.71	WMS
Total	474		

Model	Reliability	ICC	N4rel.70	N4rel.80	N4rel.90	N4rel.95
One way:						
Biased	0.791	0.177	10.83	18.57	41.78	88.21
Unbiased	0.791	0.177	10.84	18.58	41.78	88.21

NOTES: ICC = intraclass correlation. Nrel.70 is the sample size needed to obtain a reliability of 0.70; Nrel.80 is the sample size needed to obtain a reliability of 0.80; etc. This method is explained in detail in Revicki and Hays (2005).

Table B.2
Partitioning Between-Topic-Area Variance and Within-Topic-Area Variance at Time 2

Source	Degrees of Freedom	Mean square	Label for Mean Square
Between	26	23.44	BMS
Within	421	1.77	WMS
Total	447		

Model	Reliability	ICC	N4rel.70	N4rel.80	N4rel.90	N4rel.95
One way:						
Biased	0.924	0.424	3.17	5.44	12.23	25.82
Unbiased	0.924	0.424	3.17	5.44	12.23	25.82

NOTES: ICC = intraclass correlation. Nrel.70 is the sample size needed to obtain a reliability of 0.70; Nrel.80 is the sample size needed to obtain a reliability of 0.80; etc. This method is explained in detail in Revicki and Hays (2005).

Table B.3
Spearman Correlations Between Ranks and Ratings of Importance of Plan Topic Areas
at Time 1 (N = 18)

	Access	Avail	Compla	Comm	Coord	Plan Sr	Author	Ancill	Prev
Access	-0.29								
Avail		-0.64**							
Compla			-0.37						
Comm				-0.72**					
Coord					-0.19				
Plan Sr						-0.74**			
Author							-0.82**		
Ancill								-0.49*	
Prev									-0.62**

* Statistically significant at the $p < 0.05$ level; ** statistically significant at the $p < 0.01$ level.

NOTES: Access = Access to care; Avail = Provider availability; Compla = Complaints and appeals; Comm = Provider communication; Coord = Coordination of care; Plan Sr = Health plan services; Author = Authorization of care; Ancill = Ancillary clinical services; Prev = Preventive care.

Table B.4
Spearman Correlations Between Ranks and Ratings of Importance of Plan Topic Areas
at Time 2 (N = 17)

	Access	Avail	Compla	Comm	Coord	Plan Sr	Author	Ancill	Prev
Access	-0.52*								
Avail		0.00							
Compla			-0.63**						
Comm				-0.48					
Coord					-0.32				
Plan Sr						-0.63**			
Author							-0.20		
Ancill								-0.39	
Prev									-0.14

* Statistically significant at the $p < 0.05$ level; ** statistically significant at the $p < 0.01$ level.

NOTES: Access = Access to care; Avail = Provider availability; Compla = Complaints and appeals; Comm = Provider communication; Coord = Coordination of care; Plan Sr = Health plan services; Author = Authorization of care; Ancill = Ancillary clinical services; Prev = Preventive care.

Table B.5
Spearman Correlations Between Ranks of Plan Topic Areas from Time 1 to Time 2 (N = 17)

	Access	Avail	Compla	Comm	Coord	Plan Sr	Author	Ancill	Prev
Access	0.80**								
Avail		0.33							
Compla			0.60*						
Comm				0.66**					
Coord					0.67**				
Plan Sr						0.72**			
Author							0.55*		
Ancill								0.69**	
Prev									0.48*

* Statistically significant at the $p < 0.05$ level; ** statistically significant at the $p < 0.01$ level.

NOTES: Access = Access to care; Avail = Provider availability; Compla = Complaints and appeals; Comm = Provider communication; Coord = Coordination of care; Plan Sr = Health plan services; Author = Authorization of care; Ancill = Ancillary clinical services; Prev = Preventive care.

Table B.6
Spearman Correlations Between Ratings of Plan Topic Areas from Time 1 to Time 2 (N = 17)

	Access	Avail	Compla	Comm	Coord	Plan Sr	Author	Ancill	Prev
Access	0.36								
Avail		0.28							
Compla			0.43						
Comm				0.84**					
Coord					0.43				
Plan Sr						0.83**			
Author							0.62*		
Ancill								0.73**	
Prev									0.81**

*Correlation is statistically significant at the $p < 0.05$ level; ** correlation is statistically significant at the $p < 0.01$ level.

NOTES: Access = Access to care; Avail = Provider availability; Compla = Complaints and appeals; Comm = Provider communication; Coord = Coordination of care; Plan Sr = Health plan services; Author = Authorization of care; Ancill = Ancillary clinical services; Prev = Preventive care.

APPENDIX C

NEW CAHPS SUPPLEMENTAL ITEMS FOR QUALITY IMPROVEMENT

Getting health care from a specialist

- Q9A. What was the main reason you had trouble seeing a specialist?
- Q20A. In the last 12 months, how much of a problem, if any, was it to wait as long as you did to see a provider?
- Q20B. In the last 12 months, how often did you have a problem getting an appointment because the provider you wanted to see worked limited hours or had few available appointments?
- Q34A. In the last 12 months, did you get care from more than one kind of health care provider or use more than one kind of health care service?
- 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?
- Q34C. In the last 12 months, did anyone from your health plan, doctor's office, or clinic help coordinate your care among the different health care providers or services?
- Q34D. In the last 12 months, who helped to coordinate your care? Mark one or more.
- Q34E. How satisfied are you with the help you received to coordinate your care in the last 12 months?

After-hours care

- A. 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?
- B. In the last 12 months, how much of a problem, if any, was it to get the after-hours care you needed?
- C. What was the main reason you had a problem getting after-hours care?

After-hours telephone contact

- 5A. In the last 12 months, did you call your personal doctor or nurse after regular office hours?
- 5B. In the last 12 months, how much of a problem, if any, was it to talk to your personal doctor or nurse by telephone after regular office hours?**
- 5C. What was the main reason you had a problem talking to your personal doctor or nurse by telephone after regular office hours?

Your Health Plan–Information

- 40A. What kind of information did you have a problem understanding or using? Mark one or more.**
- 40B. Where did you get the information you had a problem understanding or using? Mark one or more.
- 40C. When you looked for information in the last 12 months, did you go to your health plan's Internet site?
- 40D. How useful was the information you found on your health plan's Internet site?
- 40E. In the last 12 months, did you use information on your health plan's Internet site to help choose a health care provider?

Your Health Plan–Customer service

- 41A. In the last 12 months, how often did your health plan's customer service staff treat you with courtesy and respect?
- 41B. 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?
- 42A. What was the main reason you had a problem getting help from your health plan's customer service?**
- 42B. How many calls did it take for you to get the help you needed from your health plan's customer service?

Existing CAHPS Items That Were Modified:

20. In the last 12 months, not counting the times you needed health care right away, how many days did you usually have to wait between making an appointment and actually seeing a provider? (This item had a modified set of response categories.)
- 40. In the last 12 months, how much of a problem, if any, was it to understand or use this information? (This item had wording modified from the previous existing CAHPS item.)**

Item Type Information:

Categorical Variables: 9A, 34D, C, 5C, 40A, 40B, 42A

Ordinal Variables: 20 (existing), 20A, 20B, 34B, 34E, B, 5B, 40 (existing), 40D, 41A, 41B, 42B

Dichotomous variables: 34A, 34C, A, 5A, 40C, 40E

Categorical means according to or using categories. Dichotomous means having only two values. Ordinal means having variable classifications that have ordered values.

APPENDIX D

SUPPLEMENTAL QUALITY IMPROVEMENT (QI) ITEMS IN THE FIELD TEST

This document contains the full set of QI items proposed for testing. The QI items are labeled with letters after the number for the corresponding CAHPS item for which more information is being sought. The item set has been updated to reflect the comments we received and discussed. The items address the following topic areas:

- Access to specialist care
- Access to routine care
- Access to after-hours care
- Communication across providers
- Coordination of care across providers and services
- Customer service
- Health plan information and materials.

Existing CAHPS 3.0 items are included to provide a preview of item placement and to set item context. There are a total of 24 QI items in this document.

Highlighted in **RED** and bolded below are the actual percentage of responses for an item that has:

1. A very low percentage of responses (defined as 4 percent or below for a response option OR 18 percent or lower for a yes/no item) OR
2. A very large percentage of responses in the “other” reason category OR over 50 percent of responses in any given reason category.

Getting Health Care From A Specialist
--

Existing CAHPS 3.0 Item

9. In the last 12 months, how much of a problem, if any, was it to see a specialist that you needed to see?

A big problem → **Go to**

Question 9A

A small problem → **Go to**

Question 9A

Not a problem

Proposed addition to 3.0 to obtain more focused and specific data on Access to Specialist Care (new item)

9A. What was the main reason you had a problem seeing a specialist?

My doctor did not think I needed to see a specialist

My health plan approval or authorization was delayed or denied

I wasn't sure where to find a list of specialists in my health plan or network **(3.2 percent)**

The specialists I had to choose from were too far away

I did not have enough specialists to choose from

The specialist I wanted did not belong to my health plan or network

I could not get an appointment at a time that was convenient

Some other reason **(22 percent)**

Existing CAHPS 3.0 Item MODIFIED

20. In the last 12 months, not counting the times you needed health care right away, how many days did you usually have to wait between making an appointment and actually seeing a provider?

- Same day
- 1 day
- 2-3 days
- 4-7 days
- 8-14 days
- 15-30 days
- 31-60 days
- 61-90 days
- 91 days or longer

Proposed addition to 3.0 to obtain more focused and specific data on Access to Routine Care (new item)

20A. In the last 12 months, how much of a problem, if any, was it to wait as long as you did to see a provider?

- A big problem → Go to

Question 20B

- A small problem → Go to

Question 20B

- Not a problem

Optional item to supplement 20A (new item)

20B. In the last 12 months, how often did you have a problem getting an appointment because the provider you wanted to see worked limited hours or had few available appointments?

- Never
- Sometimes
- Usually
- Always

~~~~~  
**Existing CAHPS 3.0 Item**

**34. Using any number from 0 to 10, where 0 is the worst health care possible and 10 is the best health care possible, what number would you use to rate all your health care?**

- 0 Worst health care possible
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10 Best health care possible

*Optional item to obtain more focused and specific data on Coordination of Care Across Providers and Services (adapted from CAHPS 3.0 child commercial core)*

**34A.** In the last 12 months, did you get care from more than one kind of health care provider or use more than one kind of health care service?

Yes → Go to Question 34B

No → Go to Question 35

*Optional item to obtain more focused and specific data on Communication Across Providers and (adapted from Group-CAHPS 3.0)*

**34B.** 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?

Never

Sometimes

Usually

Always

*Proposed addition to 3.0 to obtain more focused and specific data on Coordination of Care Across Providers and Services (adapted from CAHPS 3.0 child commercial core)*

**34C.** In the last 12 months, did anyone from your health plan, doctor's office or clinic help coordinate your care among the different health care providers or services?

Yes → Go to Question 34D

No → Go to Question 35

*Optional item to supplement 34C (new item)*

**34D.** In the last 12 months, who helped to coordinate your care? Mark one or more.

Someone from my health plan

Someone from my doctor's office or clinic

Someone from another organization

A friend or family member

Me

*Optional item to supplement 34C (new item)*

**34E.** How satisfied are you with the help you received to coordinate your care in the last 12 months?

Very dissatisfied

Dissatisfied

Neither dissatisfied nor satisfied

Satisfied

Very satisfied

After-Hours Care

Items A–C of this new section would be inserted in CAHPS 3.0 between the existing sections “Your Health Care In The Last 12 Months” and “Your Health Plan.”

**Proposed addition to 3.0 to obtain more focused and specific data on After-Hours Care (adapted from G-CAHPS 3.0 supplemental item)**

**A. 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?**

Yes → **Go to Question B**

Total N=88 respondents (12 percent)

No → **Go to Question 35**

**Proposed addition to 3.0 to obtain more focused and specific data on After-Hours Care (new item)**

**B. In the last 12 months, how much of a problem, if any, was it to get the after-hours care you needed?**

A big problem → **Go to Question C**

A small problem → **Go to Question C**

Not a problem → **Go to Question 35**

**Optional item to supplement A and B (new item)**

**C. What was the main reason you had a problem getting after hours care?**

I did not know where to go for after-hours care.

I wasn’t sure where to find a list of doctors’ offices or clinics in my health plan or network that are open for after-hours care.

The doctor’s office or clinic that had after-hours care was too far away.

The doctor’s office or clinic that had after-hours care was closed.

The only place I could get after-hours care was the Emergency Room (ER). (50 percent)

Some other reason

**Optional item to obtain more focused and specific data on after-hours telephone contact (new item). Items 5A, 5B, and 5C would be inserted in “Your Personal Doctor or Nurse” section.**

**5A. In the last 12 months, did you call your personal doctor or nurse after regular office hours?**

Yes → **Go to Question 5B**  
(17.5 percent)

No → **Go to Question 6**

*Optional item to obtain more focused and specific data on after-hours telephone contact (new item).*

**5B. In the last 12 months, how much of a problem, if any, was it to talk to your personal doctor or nurse by telephone after regular office hours?**

- A big problem
- A small problem
- Not a problem

*Optional item to obtain more focused and specific data on after-hours telephone contact (new item).*

**5C. What was the main reason you had a problem talking to your personal doctor or nurse by telephone after regular office hours?**

- I did not know what number to call.
- I left a message, but no one returned my call.
- I could not leave a message at the number I called.
- Another provider was covering for my personal doctor or nurse.
- Some other reason (47 percent)

Your Health Plan

**Existing CAHPS 3.0 Item**

**39. In the last 12 months, did you look for any information about how your health plan works in written materials or on the Internet?**

- Yes → **Go to Question 40**
- No → **Go to Question 41**

**Existing CAHPS 3.0 Item MODIFIED**

**40. In the last 12 months, how much of a problem, if any, was it to understand or use this information?**

- A big problem → **Go to Question 40A**
- A small problem → **Go to Question 40A**
- Not a problem → **Go to Question 40C**

*Proposed addition to 3.0 to obtain more focused and specific data on Information and Materials (new item)*

**40A. What kind of information did you have a problem understanding or using? Mark one or more.**

- Benefits and coverage for doctor or specialist visits
- Benefits and coverage for pharmacy
- Getting a referral to a specialist
- After-hours or urgent care
- Choosing a health provider
- Getting care outside your network
- Something else

*Optional item to obtain more focused and specific data on Information and Materials (new item)*

**40B. Where did you get the information you had a problem understanding or using? Mark one or more.**

- From my employer
- From my health plan's Internet site
- From my health plan's call center
- Written materials from my health plan
- From my health plan's sales representative (3.96 percent)
- From my doctor's office
- From my insurance agent or broker (2.97 percent)
- Some other source
- Not sure where I got it

*Optional item to supplement 40A (new item)*

**40C. When you looked for information in the last 12 months did you go to your health plan's Internet site?**

- Yes → Go to Question 40D
- No → Go to Question 41

*Optional item to supplement 40A (new item)*

**40D. How useful was the information you found on your health plan's Internet site?**

- Not at all useful
- A little useful
- Somewhat useful
- Very useful

*Optional item to supplement 40A (new item)*

**40E. In the last 12 months, did you use information on your health plan's Internet site to help choose a health care provider?**

- Yes
- No



**Existing CAHPS 3.0 Item**

**41. In the last 12 months, did you call your health plan's customer service to get information or help?**

- Yes → Go to Question 41A
- No → Go to Question 43

*Optional item to obtain more focused and specific data on Customer Service (modified CAHPS 3.0 item)*

**41A. In the last 12 months, how often did your health plan's customer service staff treat you with courtesy and respect?**

- Never
- Sometimes
- Usually
- Always

*Optional item to obtain more focused and specific data on Customer Service (modified CAHPS 3.0 item)*

**41B. 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?**

- Never
- Sometimes
- Usually
- Always
- I got all the help I needed during my call

**Existing CAHPS 3.0 Item**

**42. In the last 12 months, how much of a problem, if any, was it to get the help you needed when you called your health plan's customer service?**

- A big problem → **Go to**

**Question 42A**

- A small problem → **Go to**

**Question 42A**

- Not a problem → **Go to**

**Question 42B**

*Proposed addition to 3.0 to obtain more focused and specific data on Customer Service (new item)*

**42A. What was the main reason you had a problem getting help from your health plan's customer service?**

- I had to stay on the line too long, waiting for customer service
- I could not leave a message at the number I called (1.47 percent)
- Customer service was closed when I called (3.71 percent)
- The information customer service gave me was not correct
- Customer service did not have the information I needed
- I had to wait too long for someone to call me back (2.95 percent)
- No one called me back (2.29 percent)
- Some other reason (25 percent)

*Proposed addition to 3.0 to obtain more focused and specific data on Customer Service (new item)*

**42B. How many calls did it take for you to get the help you needed from your health plan's customer service?**

- 1
- 2
- 3
- 4
- 5 or more
- I am still waiting for help



## APPENDIX E: CORRELATIONS OF SUPPLEMENTAL QI ITEMS

**Table E.1 Correlations of Supplemental QI Items**

|      | 9A | A | B | C | 5A   | 5B   | 5C | 20A  | 20B   | 40A | 40B | 40C   | 40D   | 40E   | 41A   | 41B   | 42A | 42B   | 34A   | 34B   | 34C   | 34D   | 34E   |
|------|----|---|---|---|------|------|----|------|-------|-----|-----|-------|-------|-------|-------|-------|-----|-------|-------|-------|-------|-------|-------|
| Q9A  |    |   |   |   |      |      |    |      |       |     |     |       |       |       |       |       |     |       |       |       |       |       |       |
| QA   |    |   |   |   | 0.18 | 0.11 |    | 0.14 | −0.02 |     |     | 0.06  | 0.07  | −0.13 | 0.03  | 0.10  |     | −0.08 | 0.03  | 0.12  | −0.09 |       | 0.08  |
| QB   |    |   |   |   | 0.02 | 0.31 |    | 0.30 | −0.42 |     |     | −0.14 | 0.55  | −0.03 | 0.26  | 0.08  |     | 0.29  | 0.05  | 0.26  | −0.02 |       | 0.43  |
| QC   |    |   |   |   |      |      |    |      |       |     |     |       |       |       |       |       |     |       |       |       |       |       |       |
| Q5A  |    |   |   |   |      | 0.02 |    | 0.01 | 0.16  |     |     | 0.02  | −0.00 | 0.14  | 0.01  | −0.00 |     | 0.03  | 0.03  | 0.01  | −0.00 |       | −0.05 |
| Q5B  |    |   |   |   |      |      |    | 0.31 | 0.28  |     |     | 0.05  | 0.37  | −0.08 | 0.14  | −0.01 |     | 0.24  | 0.20  | 0.40  | −0.16 |       | 0.15  |
| Q5C  |    |   |   |   |      |      |    |      |       |     |     |       |       |       |       |       |     |       |       |       |       |       |       |
| Q20A |    |   |   |   |      |      |    |      | 0.18  |     |     | −0.00 | 0.23  | 0.12  | 0.17  | 0.13  |     | −0.02 | 0.05  | 0.32  | −0.04 |       | 0.23  |
| Q20B |    |   |   |   |      |      |    |      |       |     |     | 0.25  | −0.36 | 0.01  | 0.01  | 0.04  |     | −0.13 | 0.05  | −0.39 | 0.19  |       | 0.01  |
| Q40A |    |   |   |   |      |      |    |      |       |     |     |       |       |       |       |       |     |       |       |       |       |       |       |
| Q40B |    |   |   |   |      |      |    |      |       |     |     |       |       |       |       |       |     |       |       |       |       |       |       |
| Q40C |    |   |   |   |      |      |    |      |       |     |     |       |       |       | −0.02 | −0.07 |     | −0.00 | 0.08  | −0.00 | −0.06 |       | −0.10 |
| Q40D |    |   |   |   |      |      |    |      |       |     |     |       |       | −0.17 | 0.39  | 0.30  |     | 0.20  | −0.00 | 0.15  | 0.10  |       | 0.26  |
| Q40E |    |   |   |   |      |      |    |      |       |     |     |       |       |       | 0.09  | 0.05  |     | −0.01 | 0.24  | −0.01 | 0.01  |       | 0.13  |
| Q41A |    |   |   |   |      |      |    |      |       |     |     |       |       |       |       | 0.46  |     | 0.15  | 0.02  | 0.29  | −0.03 |       | 0.20  |
| Q41B |    |   |   |   |      |      |    |      |       |     |     |       |       |       |       |       |     | 0.36  | −0.02 | 0.05  | 0.07  |       | 0.11  |
| Q42A |    |   |   |   |      |      |    |      |       |     |     |       |       |       |       |       |     |       |       |       |       |       |       |
| Q42B |    |   |   |   |      |      |    |      |       |     |     |       |       |       |       |       |     |       | −0.09 | 0.03  | 0.11  |       | 0.07  |
| Q34A |    |   |   |   |      |      |    |      |       |     |     |       |       |       |       |       |     |       |       |       |       |       |       |
| Q34B |    |   |   |   |      |      |    |      |       |     |     |       |       |       |       |       |     |       |       |       |       | −0.21 | 0.29  |
| A34C |    |   |   |   |      |      |    |      |       |     |     |       |       |       |       |       |     |       |       |       |       |       |       |
| Q34D |    |   |   |   |      |      |    |      |       |     |     |       |       |       |       |       |     |       |       |       |       |       |       |
| Q34E |    |   |   |   |      |      |    |      |       |     |     |       |       |       |       |       |     |       |       |       |       |       |       |

\* Correlation is statistically significant at the  $p < 0.05$  level; \*\* correlation is statistically significant at the  $p < 0.01$  level.

NOTES: Correlations reflect coding in which “high” is always good for both candidate QI items and composites. Results are not applicable for categorical variables, such as 9A, A, B, C, and are shaded. Appendix C includes a complete list of the topic and item labels.



## REFERENCES

- Agency for Healthcare Research and Quality (AHRQ), *CAHPS 2.0 Survey and Report Kit*, Rockville, Md.: U.S. Department of Health and Human Services, AHRQ Publication No. 99-039, 1999.
- Carman, K., P. F. Short, D. O. Farley, et al., “Epilogue: Early Lessons from CAHPS Demonstrations and Evaluations,” *Medical Care (Supplement)*, Vol. 37, No. 3, 1999, pp. MS97–MS105.
- Farley, D. O., M. E. Elliott, P. F. Short, D. E. Kanouse, R. D. Hays, “Effects of CAHPS® Health Plan Performance Information on Health Plan Choices by Iowa Medicaid Beneficiaries,” *Medical Care Research and Review*, Vol. 59, No. 3, 2002b, pp. 319–336.
- Farley, D. O., P. F. Short, M. N. Elliott, D. E. Kanouse, J. A. Brown, R. D. Hays, “Effects of CAHPS® Health Plan Performance Information on Plan Choices by New Jersey Medicaid Beneficiaries,” *Health Services Research*, Vol. 37, No. 4, 2002a, pp. 985–1008.
- Fox, M. H., J. Moore, M. Zimmerman, S. Hill, and C. H. Foster, “The Effectiveness of CAHPS Among Women Enrolling in Medicaid Managed Care,” *Journal of Ambulatory Care Management*, Vol. 24, No. 4, October 2001, pp. 76–91.
- Kanouse, D. E., M. E. Spranca, J. Uhrig, M. N. Elliott, P. F. Short, D. O. Farley, and R. D. Hays, “Effects of CAHPS Reports on Medicaid Recipients’ Choice of Health Plans in a Laboratory Setting,” Paper presented at the meeting of the Association for Health Services Research, Los Angeles, Calif., 2000.
- Linstone, Harold A., *The Delphi Method: Techniques and Applications*, Addison-Wesley Pub. Co., 1975.
- National Committee on Quality Assurance (NCQA), *HEDIS® 2002 Volume 3: Specification for Survey Measures*, Washington DC: National Committee for Quality Assurance, 2002.
- Quigley, Denise D., Dennis Scanlon, Donna O. Farley, Han de Vries, “The Utility of CAHPS for Health Plans,” Santa Monica, Calif.: RAND Corporation, WR-115-AHRQ, 2003. Available at [http://www.rand.org/pubs/working\\_papers/WR115/index.html](http://www.rand.org/pubs/working_papers/WR115/index.html).
- Revicki, D., Hays, R. D., “Reliability and Validity (Including Responsiveness),” in Peter Fayers and Ron Hays, eds., *Assessing Quality of Life in Clinical Trials*, 2nd Ed., Oxford University Press, 2005, pp. 25–39.
- Sackman, Harold, *Delphi Assessment: Expert Opinion, Forecasting, and Group Process*, Santa Monica, Calif.: RAND Corporation, R-1283, 1974. Available at <http://www.rand.org/pubs/reports/R1283>.
- Spranca, M. E., D. E. Kanouse, P. F. Short, D. O. Farley, and R. D. Hays, “Do Consumer Reports of Health Plan Quality Affect Health Plan Selection?” *Health Services Research*, Vol. 35, No. 5 (part 1), 2000, pp. 933-947.