# Measuring Beneficiary Knowledge of the Medicare Program: A Psychometric Analysis

Carla M. Bann, Ph.D., Sherry A. Terrell, Ph.D., Lauren A. McCormack, Ph.D., and Nancy D. Berkman, Ph.D.

*Reliable measures of Medicare beneficiaries' program knowledge are necessary for credible program monitoring, evaluation, and public accountability. This study developed and evaluated the psychometric properties of two possible measures of beneficiary knowledge. One measure was based on self-reported knowledge, the other was a true/false quiz which requires beneficiaries to demonstrate their knowledge. We used data from the 1998 and 1999 Medicare Current Beneficiary Survey (MCBS) to evaluate the reliability and construct validity of the indices. Overall, based on both content considerations and the psychometric analyses, the true/false quiz proved to be the more accurate and useful measure of beneficiaries' knowledge.*

## INTRODUCTION

Given the increasing variety of health insurance options available to Medicare beneficiaries, it is of vital importance that beneficiaries sufficiently understand the tradeoffs associated with their different insurance options so they can make informed choices. However, several studies have shown that Medicare beneficiaries possess a low level of understanding about their health care coverage (Gibbs, Sangl, and Burrus, 1996; Hibbard et al., 1998; Murray and Shatto, 1998). Many beneficiaries do not understand what plan options are available or what services are covered and several have never heard of a Medicare health maintenance organization (HMO). Furthermore, the majority of beneficiaries cannot identify basic distinctions between original Medicare and Medicare managed care plans (Hibbard et al., 1998).

The goal of the current study was to develop and compare possible measures of beneficiary knowledge about the Medicare Program. One useful data source is the MCBS which contains several questions to measure beneficiaries' knowledge of the Medicare Program. We developed two different knowledge indices, using questions administered during the 1998 and 1999 MCBS[1]. The first measure, the perceived knowledge index (PKI), assesses how much beneficiaries feel they know about five different aspects of the Medicare Program. The second index, a seven-item quiz, tests beneficiaries' knowledge, using a set of true/false questions about Medicare options and managed care plans.

We evaluated the psychometric properties of these new knowledge indices, including their internal consistency reliability and construct validity. We also analyzed the properties of the individual items comprising each scale. Based on these performance measures, we made recommendations about the use of these indices for assessing beneficiary knowledge of the Medicare Program.

[1] The supplemental rounds of the 1998 and 1999 MCBS included in this study were administered from January to August 1999 and 2000, respectively.

## METHODS

### Data Source

The MCBS utilizes a rotating panel design in which a large national probability sample of 14,000 or more Medicare beneficiaries is interviewed every 4 months for up to 4 years. Each year, approximately one-quarter of the sample is rotated out of the survey and replaced with new sample members. Therefore, 25 percent of each annual MCBS data set represents a cross section of the Medicare population continuously enrolled in the program from January 1st of that year; 75 percent represents a longitudinal panel of beneficiaries.

During each round, participants in the MCBS received a core questionnaire with questions on health insurance, utilization of health care services, and expenditures. In some rounds, the participants also received additional questionnaires or supplements. The winter MCBS round includes the beneficiary knowledge supplemental questionnaire; the spring round includes the beneficiary information needs supplement. CMS releases two types of public use data files for the MCBS. The Access to Care File contains data from the supplemental questionnaires, addressing topics, such as satisfaction with care; the Costs and Use File contains information on service utilization and expenditures. Data from the 1998 and 1999 access to care beneficiary knowledge (rounds 23 and 26) and beneficiary information needs (rounds 24 and 27) supplemental rounds of the MCBS were used for the psychometric analyses in this study. In 1998, all respondents received the knowledge questions. However, during 1999, respondents participating in the MCBS for the first time were the only ones to receive the seven-item quiz. Therefore, fewer respondents were available for the 1999 analyses of this quiz.

Only participants who were living in the community received the knowledge questions and were included in the analyses. In addition, because the population was elderly and likely to experience disabilities, the MCBS used a proxy, as necessary, to obtain information about a respondent. In both 1998 and 1999, proxy interviews comprised 10 percent of all interviews conducted. Because results from earlier studies indicated that proxy and sample member participants tended to respond to the knowledge indices differently (Bann et al., 2000), we analyzed these two groups separately. Results for the sample members are presented here; the analyses of proxy respondents are available in Bann and Berkman (2002). The total number of sample member respondents living in the community who completed the perceived knowledge index was 12,524 in 1998 and 12,606 in 1999; the sample sizes for the seven-item quiz were 13,062 in 1998 and 3,920 in 1999.

## MEASURES

### PKI

The PKI includes five items that were administered during rounds 24 and 27 of the MCBS. As shown in Table 1, these questions ask beneficiaries how much they feel they know about five topics related to the Medicare Program. For each question, beneficiaries rate their knowledge on a five-point scale from "almost none of what you need to know" to "just about everything you need to know." PKI scores were created by reverse coding and summing the items; higher scores reflect a higher level of self-reported knowledge. Missing values were imputed by substituting the mean of the remaining items for the missing item values (Chapman, 1976); a technique used in other scales, such as the Short Form-36®

**Table 1**

**Medicare Current Beneficiary Survey Knowledge Questions**

| Scale/Question |
| --- |

**Perceived Knowledge Index[1]**
How much do you feel you know about what medical services Medicare covers or does not cover?

How much do you feel you know about how much you have to pay for medical services?

How much do you feel you know about supplemental or Medigap insurance, such as what it covers or how it works with Medicare to pay medical claims?

How much do you feel you know about the availability and benefits of Medicare managed care plans?

How much do you feel you know about choosing or finding a doctor or other health care provider?

**Seven-Item Quiz[2]**
Most people covered by Medicare can select among different kinds of health plan options within Medicare. (True)

Medicare without a supplemental policy pays for all of your health care expenses. (False)

The Medicare program has begun to offer more information and help in order to answer your Medicare questions. (True)

People can report complaints to Medicare about their Medicare managed care plans (HMOs) or supplemental plans if they are not satisfied with them. (True).

If someone joins a Medicare managed care plan (HMO) that covers people on Medicare, they have limited choices about which doctors they can see. (True)

If someone joins a Medicare managed care plan (HMO) that covers people on Medicare, they can change or drop the plan and still be covered by Medicare. (True)

Medicare managed care plans (HMOs) that cover people on Medicare often cover more health services, like prescribed medicines, than Medicare without a supplemental policy. (True)

**Global Perceived Knowledge Question[1]**
How much do you think you know about the Medicare program?

[1] Response options for the global perceived knowledge question and the perceived knowledge index questions were: just about everything you need to know, most of what you need to know, some of what you need to know, a little of what you need to know, and almost none of what you need to know.

[2] Correct answers to quiz questions are shown in parentheses.

NOTE: HMO is health maintenance organization.

SOURCES: Centers for Medicare & Medicaid Services: Data from the 1998 and 1999 Medicare Current Beneficiary Survey, Beneficiary Knowledge Supplemental Rounds 23 and 26 and the Beneficiary Information Needs Supplemental Rounds 24 and 27.

(Ware, Snow, and Kosinski, 2000). This imputation method was selected because an individual's own responses were expected to provide a more accurate representation of his or her knowledge than those of other respondents. Imputation was used only for respondents who answered at least one-half of the items (i.e., three of the five items)[2]. Individuals missing responses to more than one-half of the items were assigned a value of missing for the index.

---

[2] Missing values were imputed for 2.6 percent of respondents in 1998 and 1.8 percent of respondents in 1999.

## Seven-Item Quiz

The seven-item quiz, administered in MCBS rounds 23 and 26, includes true/false questions about Medicare options and Medicare managed care plans (Table 1). Each item contains only one correct answer. Quiz scores were computed as the number of items a respondent answered correctly; thus, scores ranged from zero to seven with higher values indicating greater knowledge of Medicare. Don't know responses were considered to be incorrect,

an approach utilized in similar studies (Hibbard et al., 1998; McCormack et al., 2001).

## Global Perceived Knowledge Question

The MCBS includes a global perceived knowledge question that asks respondents how much they think they know about the Medicare Program (Table 1). Responses follow the standard MCBS five-point rating convention, ranging from "almost none of what you need to know" to "just about everything you need to know." For this study, we reverse coded the global perceived knowledge question so that higher values indicate more knowledge. We expected beneficiaries perceived and demonstrated knowledge scores to be related to the level of knowledge indicated by the global knowledge question and used it as a criterion measure.

## PSYCHOMETRIC ANALYSIS METHODS

To evaluate the psychometric properties of the knowledge measures, both item- and scale-level analyses were conducted. The item-level analyses examined the psychometric properties of each individual item included in the scale; the scale-level analyses evaluated the reliability and validity of the composite scale as a whole.

### Item-Level Analyses

For both knowledge indices, we computed the correlations between each item on the index and the total index score. Item-total score correlations provide information about the contribution of the item to the reliability of the scale. Ideally, items should have item-total correlations of at least 0.30; usually, items with low correlations are removed from the scale. To avoid

possible inflation due to inclusion of the item itself in the calculation of the index score, we corrected the item-total score correlations in this article for overlap (Howard and Forehand, 1962).

In addition, because the seven-item quiz questions may be scored as correct or incorrect, we used three different approaches to evaluate the difficulty levels of the questions. First, we calculated the percentage of beneficiaries answering each item correctly. Second, we used item response theory (IRT) analyses to evaluate the properties of the items. IRT uses a model to describe the relationship between an individual's response to an item and the underlying construct (i.e., knowledge). Third, we examined the highest school grade completed by each beneficiary. Then we matched the level of difficulty for each item to a grade level. If possible, a knowledge index should contain items with a wide range of difficulty levels to enable it to discriminate among respondents who have a variety of knowledge.

### Scale-Level Analyses

For each scale, we computed the mean and standard deviation (SD) of the scores to determine the most representative scale scores and to examine the variability of scores. A lack of variability can compromise the validity of scale scores. In addition, we evaluated the reliability and construct validity of the knowledge indices.

#### Reliability

We estimated the internal consistency reliability of the scales, using Cronbach's alpha coefficient (Cronbach, 1951). Internal consistency measures the degree to which items on a scale are related to each other and therefore, appear to measure the same construct. For the index score to be reliable

for use in group-level statistical analyses, a common rule of thumb is to require alpha coefficients to be 0.70 and above (Guilford, 1956; Nunnally, 1978).

Factor Analyses

We used factor analysis to evaluate the dimensionality of the knowledge measures. For both scales, we hypothesized that the items formed a single, unidimensional construct representing knowledge of the Medicare Program. To test this hypothesis, confirmatory factor analyses, specifying one factor, were fit to each scale. We estimated the models using LISREL 8 software (Jöreskog and Sörbom, 1996). Because the items were categorical rather than continuous, we first computed polychoric correlations and the corresponding asymptotic covariance matrix for the items and used these matrices as input for the confirmatory factor analyses.

Validity

Several different approaches may establish the construct validity of a new scale. Ideally, the new scale would be shown to be highly related to a well-established and validated scale measuring the same construct, often called a "gold standard." Unfortunately, a gold standard for measuring Medicare beneficiary knowledge does not exist. Instead, we used three alternative approaches to assess the validity of the indices. Taken individually, these methods may not provide evidence as strong as that of a comparison to a gold standard; however, combined, they provide useful evidence to support the validity of the knowledge indices.

The first set of validity analyses examined the relationship between the knowledge indices and another measure of knowledge, the global perceived knowledge question. We expected beneficiaries' self-perceived knowledge to be positively related to the level of knowledge indicated by their scores on the indices. A strong relationship between this question and the knowledge indices would support the construct validity of the indices.

Next, we determined if the knowledge scale scores discriminated among groups of Medicare beneficiaries who previously showed differences in their knowledge of the Medicare Program. This approach is sometimes referred to as known-groups comparisons. Prior research served as the basis for our expectation of differing levels of knowledge among beneficiaries. For example, factors related to socioeconomic status are often predictive of levels of insurance knowledge. Several studies report that respondents with more education have higher levels of insurance knowledge (Lambert, 1980; Marquis, 1983; McCall, Rice, and Sangl, 1986; Hibbard et al., 1998; McCormack et al., 2002). Higher knowledge levels have also been associated with higher incomes (Lambert, 1980; Marquis, 1983; McCall, Rice, and Sangl, 1986; Hibbard et al., 1998) and having a supplemental insurance plan (Cafferata, 1984). Other researchers have found that, among older adults, those who are younger have more insurance-related knowledge (Lambert, 1980; Cafferata, 1984).

Based on this research, we expected that the following groups of beneficiaries would have higher levels of knowledge about the Medicare Program: (1) beneficiaries with more education, (2) beneficiaries with higher incomes, and (3) beneficiaries with supplemental insurance. We also expected that among beneficiaries who are eligible for Medicare because of their age, those who are younger (i.e., beneficiaries age 65 to 75) would have more

**Table 2**

**Means (and Standard Deviations) of Perceived Knowledge Index Questions[1]: 1998 and 1999**

| Item | 1998 | 1999 |
|---|---|---|
| Services Medicare Covers | 3.6 (1.2) | 3.7 (1.2) |
| Paying for Medical Services | 3.1 (1.2) | 3.2 (1.2) |
| Supplemental Insurance | 2.4 (1.4) | 2.4 (1.5) |
| Medicare Managed Care Plans | 3.2 (1.3) | 3.3 (1.3) |
| Choosing a Doctor | 2.9 (1.4) | 2.9 (1.4) |

[1] Scores for these items ranged from 1 to 5.

SOURCE: Centers for Medicare & Medicaid Services: Data from the Medicare Current Beneficiary Survey, Beneficiary Information Needs Supplemental Rounds 24 and 27.

program knowledge. In this study, a scale will have demonstrated construct validity if the results of the analyses on that scale showed these expected patterns.

We also compared knowledge scores for groups of respondents that we hypothesized would differ in knowledge. We expected that beneficiaries enrolled in managed care during the past year might have more program knowledge than those not enrolled in managed care during the same period. The former must make more choices regarding their insurance arrangements than beneficiaries using fee-for-service. Because four items on the seven-item quiz concern managed care, the greatest effect should be present for this measure. Finally, we thought that beneficiaries who have more experience with the Medicare Program might have higher levels of knowledge. For example, Cafferata (1984) found that, among a subsample of older adults with private insurance, service utilization was positively associated with knowledge.

For the validity analyses, level of service utilization served as an approximation of experience with the Medicare Program. We included two types of service utilization during the past year: (1) institutional utilization and (2) Part B utilization. Institutional utilization includes hospice, home health agency or skilled nursing facilities care,

and inpatient or outpatient hospital visits. We also used amounts of allowable and reimbursed charges as indicators of experience with the Medicare Program. Complete information on service utilization was not available for respondents in an HMO; therefore, our analyses of these variables included only those individuals who were not enrolled in managed care during the year before the survey data were collected.

## RESULTS

This section describes the results of the psychometric analyses of the two knowledge measures, the perceived knowledge index and the seven-item quiz.

## PKI

### Item-Level Analyses

The means and SDs for the PKI items appear in Table 2. The values for the SDs indicate good variability in responses, suggesting that respondents were using all of the response categories. The highest means were for the item on services Medicare covers, indicating that respondents felt they knew the most about this topic. The item on supplemental insurance had the lowest means, indicating that, on average, respondents knew the least about this topic.

The PKI item-total score correlations for 1998 and 1999 were very similar. All of the correlations were 0.5 or greater, suggesting that these items are highly related; the items also appear to measure the same construct (i.e., self-perceived knowledge). Those items which addressed the services Medicare covers and paying for medical services were the most highly related to the underlying construct, with item–total

score correlations around 0.70. The question on supplemental insurance had correlations of 0.62 in 1998 and 0.63 in 1998; correlations for the question on choosing a doctor had a value of 0.56 in both 1998 and 1999. The item about Medicare HMOs, least related to the construct, had correlations of approximately 0.50; however, this item still contributes to the internal consistency reliability of the index.

## Scale-Level Analyses

In 1998, the mean PKI score was 15.2 with an SD of 5.0. The mean score was slightly higher in 1999, with a value of 15.6; the SD was 5.0. The values of the SDs show good variability in scores, suggesting that the index is able to discriminate among respondents with differing levels of knowledge.

Cronbach's alphas were computed to assess the internal consistency reliability of the PKI. The values of Cronbach's alpha were 0.82 in both 1998 and 1999, indicating that the PKI demonstrated strong internal consistency reliability.

## Factor Analyses

A confirmatory factor analysis model with one factor was fit to the PKI items in each year. To assess the fit of the models, we examined several fit indices, including the Goodness-of-Fit Index (GFI), Adjusted Goodness-of-Fit Index (AGFI), Comparative Fit Index (CFI), and Non-Normed Fit Index (NNFI). For these indices, values of 0.90 or higher indicate a good fit, with 1.00 being the highest possible value (Hoyle and Panter, 1995; Schumacker and Lomax, 1996). In both 1998 and 1999, the values for all of these indices were equal to 1.00. In addition, we examined the standardized root mean square residual (SRMR) for which values of 0.08 or lower indicate a good fit (Hu and Bentler, 1999). Consistent with the other fit indices, the SRMR supports a one-factor solution; the SRMR was equal to 0.009 in 1998 and 0.010 in 1999.

## Validity

### Relationship with Global Self-Perceived Knowledge

Analyses of variance (ANOVAs) indicated that PKI scores varied significantly across levels of global self-perceived knowledge for participants in both 1998 ($F(4, 12336) = 762.71$, $p < 0.0001$) and 1999 ($F(4, 3617) = 250.33$, $p < 0.0001$). To help interpret the results, we computed the mean PKI scores for each of the five response categories included in the global self-perceived knowledge question (Figure 1). Overall, the pattern of means indicates a clear, increasing relationship between the two measures. Individuals who rated their knowledge higher on the global self-perceived knowledge question received higher PKI scores.
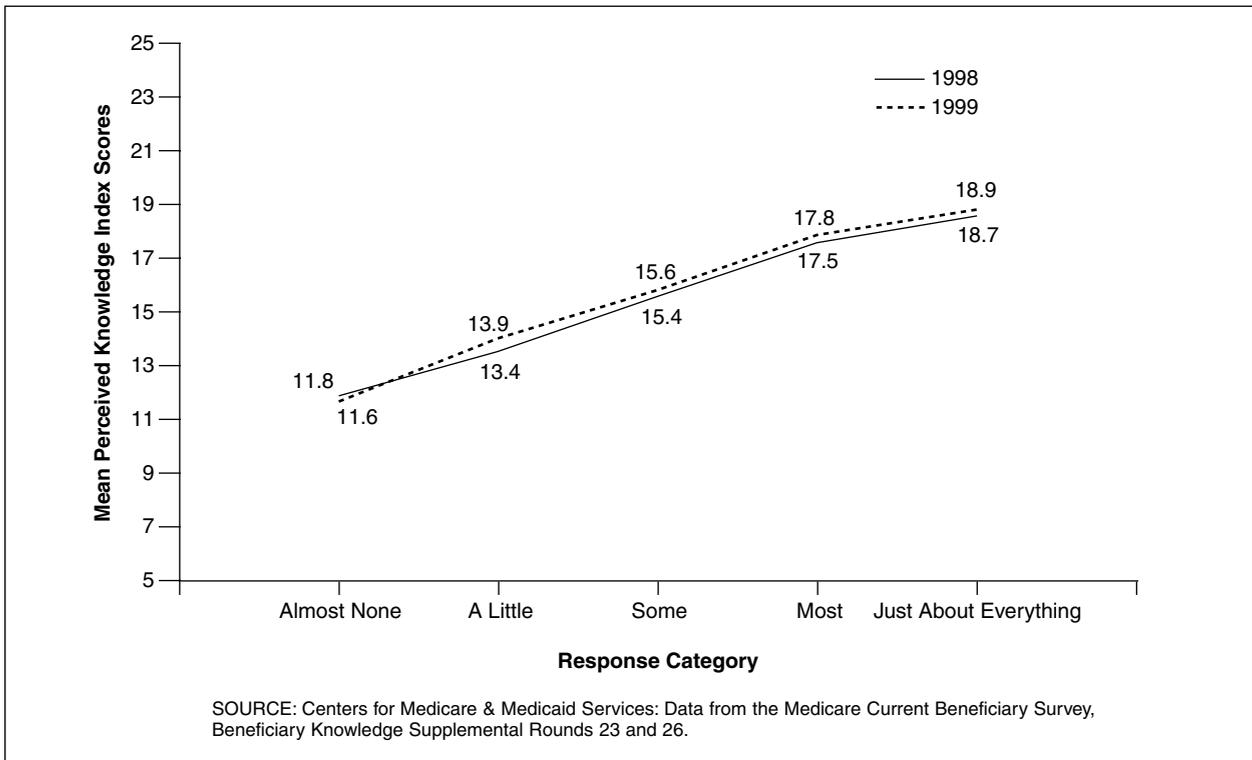
### Group Comparisons

For the next set of validity analyses, we used ANOVAs and $t$-tests to compare the PKI scores of respondents, according to various background and experience variables; these results appear in Table 3. For the $t$-tests, we tested the assumption of equal variance for the two groups, using the $F'$ (folded) statistic (Steel and Torrie, 1980). If this test was significant (i.e., the variances for the two groups were not equal), we computed an approximate $t$ statistic, using Satterthwaite's (1946) approximation to estimate the degrees of freedom.

As shown in Table 3, in both years, PKI scores differed significantly for all of the subgroup variables. An examination of the

**Figure 1**

**Mean Perceived Knowledge Index Scores, by Responses to Global Perceived Knowledge Question: 1998 and 1999**



SOURCE: Centers for Medicare & Medicaid Services: Data from the Medicare Current Beneficiary Survey, Beneficiary Knowledge Supplemental Rounds 23 and 26.

patterns of mean values indicates that the results are generally in the expected direction. Participants with more education and higher incomes received higher knowledge index scores. Also, we found higher scores for sample members between age 65 and 75, with some institutional utilization, some Part B utilization, private supplemental insurance, or enrollment in managed care. However, the means of the charges' variables did not always follow an entirely increasing pattern, possibly due to the number and range of cut-off points chosen. In some cases, the mean knowledge index scores for respondents with the most charges (e.g., $5,000 or more) were similar or smaller than the mean scores for respondents with fewer charges in the adjacent category (e.g., $500-$4,999). However, overall, the general pattern was that individuals with no charges received lower knowledge scores than those with any charges.

## Seven-Item Quiz

### Item-Level Analyses

The item-total correlations for the seven-item quiz questions are in Table 4. All of the correlations were larger than 0.30, suggesting that they are contributing to the internal consistency reliability of the scale.

We also computed the percentages of correct responses for each of the seven-item quiz questions, displayed in Table 5. We used data only from 1999 for these analyses. As previously mentioned, only participants who were new to the MCBS received the quiz questions during 1999; in prior years, all respondents received the questions. Because the participants in 1999 had not seen the quiz questions before, their responses should provide a more accurate measure of the true difficulty level of the items. As shown in Table 5, the

**Table 3**

**Subgroup Comparisons for Perceived Knowledge Index Scores: 1998 and 1999**

| Variable | 1998 | 1999 |
|---|---|---|
| Education | $F(2,12521) = 427.93$ | $F(2,12569) = 488.15$ |
| Age Category | $t(10671) = 8.09$ | $t(11240) = 13.24$ |
| Income Category | $t(8164) = -25.81$ | $t(8886) = -26.39$ |
| Managed Care | $t(3552) = -12.64$ | $t(12592) = 22.74$ |
| Private Supplemental Insurance | $t(11715) = -15.92$ | $t(1911) = -6.50$ |
| Institutional Utilization[1] | $t(10195) = -3.39$ | $t(10126) = -6.73$ |
| Part B Utilization[1] | $t(10195) = -12.29$ | $t(10126) = -11.79$ |
| Total Covered Charges[1] | $F(3,10193) = 49.37$ | $F(3,10124) = 50.72$ |
| Covered Institutional Charges[1] | $F(3,10193) = 3.83$ | $F(3,10124) = 15.78$ |
| Allowed Part B Charges[1] | $F(3,10193) = 61.91$ | $F(3,10124) = 59.03$ |

[1] Respondents enrolled in managed care during the past year were excluded from these analyses.

NOTE: All tests were significant at the $p<0.001$ level.

SOURCE: Centers for Medicare & Medicaid Services: Data from the Medicare Current Beneficiary Survey, Beneficiary Information Needs Supplemental Rounds 24 and 27.

**Table 4**

**Item-Total Correlations for the Seven-Item Quiz: 1998 and 1999**

| Quiz Item | 1998 | 1999 |
|---|---|---|
| Can select different health plan options? | 0.44 | 0.40 |
| Medicare alone pays for all health care expenses? | 0.35 | 0.38 |
| Medicare offers more information? | 0.37 | 0.37 |
| Can report complaints to Medicare about HMOs and supplemental insurance? | 0.46 | 0.43 |
| Limited choices of doctors if on HMOs? | 0.52 | 0.49 |
| Can drop HMO and still be covered by Medicare? | 0.55 | 0.53 |
| HMOs cover more health services? | 0.49 | 0.46 |

NOTE: HMO is health maintenance organization.

SOURCE: Centers for Medicare & Medicaid Services: Data from the Medicare Current Beneficiary Survey, Beneficiary Knowledge Supplemental Rounds 23 and 26.

**Table 5**

**Percentage of Correct Responses and Item Response Theory Parameters for the Seven-Item Quiz: 1999**

| Quiz Item/Grade Level | Response Rate | | Two-Parameter Logistic Model | |
|---|---|---|---|---|
| | Percent Correct | Don't Know | Slope (a) | Threshold (b) |
| Can select different health plan options? | 44 | 44 | 1.14 | 0.26 |
| College | 51 | — | — | — |
| Medicare alone pays for all health care expenses? | 78 | 17 | 1.33 | -1.27 |
| 8th Grade or Less | 65 | — | — | — |
| Medicare offers more information? | 52 | 39 | 0.99 | -0.08 |
| High School Graduate | 55 | — | — | — |
| Can report complaints to Medicare about HMOs and supplemental insurance? | 56 | 35 | 1.23 | -0.26 |
| 9th to 12th Grade, No High School Diploma | 52 | — | — | — |
| Limited choices of doctors if on HMOs? | 61 | 31 | 1.89 | -0.31 |
| 9th to 12th Grade, No High School Diploma | 59 | — | — | — |
| Can drop HMO and still be covered by Medicare? | 46 | 46 | 2.56 | 0.16 |
| College | 56 | — | — | — |
| HMOs cover more health services? | 37 | 50 | 2.14 | 0.46 |
| College | 46 | — | — | — |

NOTE: HMO is health maintenance organization.

SOURCES: Centers for Medicare & Medicaid Services: Data from the Medicare Current Beneficiary Survey, Beneficiary Knowledge Supplemental Rounds 23 and 26.

question that asks whether HMOs cover more health services was the most difficult item, with 37 percent of respondents answering this item correctly. The question about whether Medicare alone pays for all health care expenses received the highest percentage of correct responses (78 percent). This question was the easiest

item for respondents, possibly because it addresses a particularly relevant topic for beneficiaries.

In addition to computing classical test theory analyses (i.e., item-total correlations and percentages correct), we conducted item response theory analyses of the seven-item quiz questions. Results from a previous study (Bann, 2001) indicated that a two-parameter logistic (2PL) IRT model was the most appropriate model for these items. A 2PL model estimates a slope (a) parameter and a threshold (b) parameter for each item. The slope quantifies how related the item is to the construct being measured by the scale and is analogous to the item-total correlation. The threshold parameter is an indicator of the difficulty of the item; items with larger thresholds are more difficult.

The 2PL IRT parameters appear in Table 5. The IRT results corresponded exactly to those found by examining the percentage correct and item-total correlations. The item concerning whether HMOs cover more services remained the most difficult item; the item about whether Medicare alone pays all expenses was the easiest item. Also, the item about whether a beneficiary can drop an HMO and still be covered by Medicare had the highest slope; therefore, this item was the most related to the underlying construct.

Finally, to make the difficulty level of the items more meaningful, we attempted to match a grade level to the items, using information about beneficiaries' educational achievement. First, sample members in 1999 were classified into the following five categories, based on the highest school grade they completed: (1) no formal education, (2) 8th grade or less, (3) 9th to 12th grade, without a high school diploma, (4) high school graduate, and (5) college. Then we calculated the percentage of sample members answering an item correctly, according to their educational achievement.

Often, knowledge items are assigned a difficulty level based on the level at which 50 percent of respondents answered the item correctly. A consideration for true/false items or other items with only two alternatives is that respondents have a 50 percent chance of guessing the correct answer. However, several properties of the MCBS quiz questions suggest that the level of guessing should be minimal. First of all, the quiz questions actually contain three response options (true, false, don't know), rather than just two. The "don't know" option allows respondents to indicate that they do not know the answer to the question. Also, in contrast to an educational testing environment, there is not a clear incentive for guessing the answer.

To determine whether or not respondents appear to be guessing, we computed the number of "don't know" responses for each item; these values appear in Table 5. The percentages of don't know responses ranged from 17 to 50 percent, with more difficult items tending to have higher don't know percentages. These percentages suggest that a substantial number of respondents chose to indicate that they did not know the answer to the item rather than guessing. To further investigate the possibility of guessing, we evaluated these items, using a three-parameter logistic (3PL) IRT model. In addition to the slope (a) and threshold (b) parameters shown in Table 5, the model estimates a guessing (c) parameter. The guessing parameters for the items ranged from 0.01 to 0.04. These values are very close to zero, indicating a very low likelihood of guessing.

Given the low likelihood of guessing, we assigned items to a particular grade level when at least 50 percent of respondents at that grade level (as well as at all higher grade levels) answered the item correctly. Table 5 contains a list of the quiz items, their corresponding grade level, and the

percentage of respondents at the grade level who answered the item correctly. For example, 51 percent of respondents with a college education answered the question on selecting different health plan options correctly, suggesting that this item is at the college level of difficulty. As shown in the table, the seven-item quiz contains items covering all education levels, except for "no formal education." The item concerning whether Medicare alone pays for all expenses was easiest; 65 percent of those having an 8th grade education or less answered it correctly. In contrast, the item about whether HMOs cover more health services was very difficult; only 46 percent of respondents who attended college answered it correctly.

Finally, to assign a grade level to the overall seven-item quiz, at least 50 percent of beneficiaries at the grade level must have answered at least 50 percent of the questions correctly (i.e., received a quiz score of 4 or higher). The results indicate that the seven-item quiz corresponds to a high school level of difficulty. Fifty-one percent of sample members with some high school education received an overall quiz score of 4 or higher.

## Scale-Level Analyses

The average seven-item quiz scores (and SDs) were 3.76 (2.12) in 1998 and 3.74 (2.08) in 1999. The alpha coefficient for the seven-item quiz was 0.74 in 1998 and 0.73 in 1999, indicating that the seven-item quiz demonstrated good internal consistency reliability.

## Factor Analyses

The fit indices for a one-factor model of the seven-item quiz questions provide evidence that the items form a single underlying factor. The SRMR was equal to 0.067 in 1998 and 0.070 in 1999. The other fit indices were equal to 0.99, except for the AGFI in 1998 which was equal to 0.98. This finding indicates that the items may be combined into an overall score, representing knowledge of the Medicare Program.

## Validity

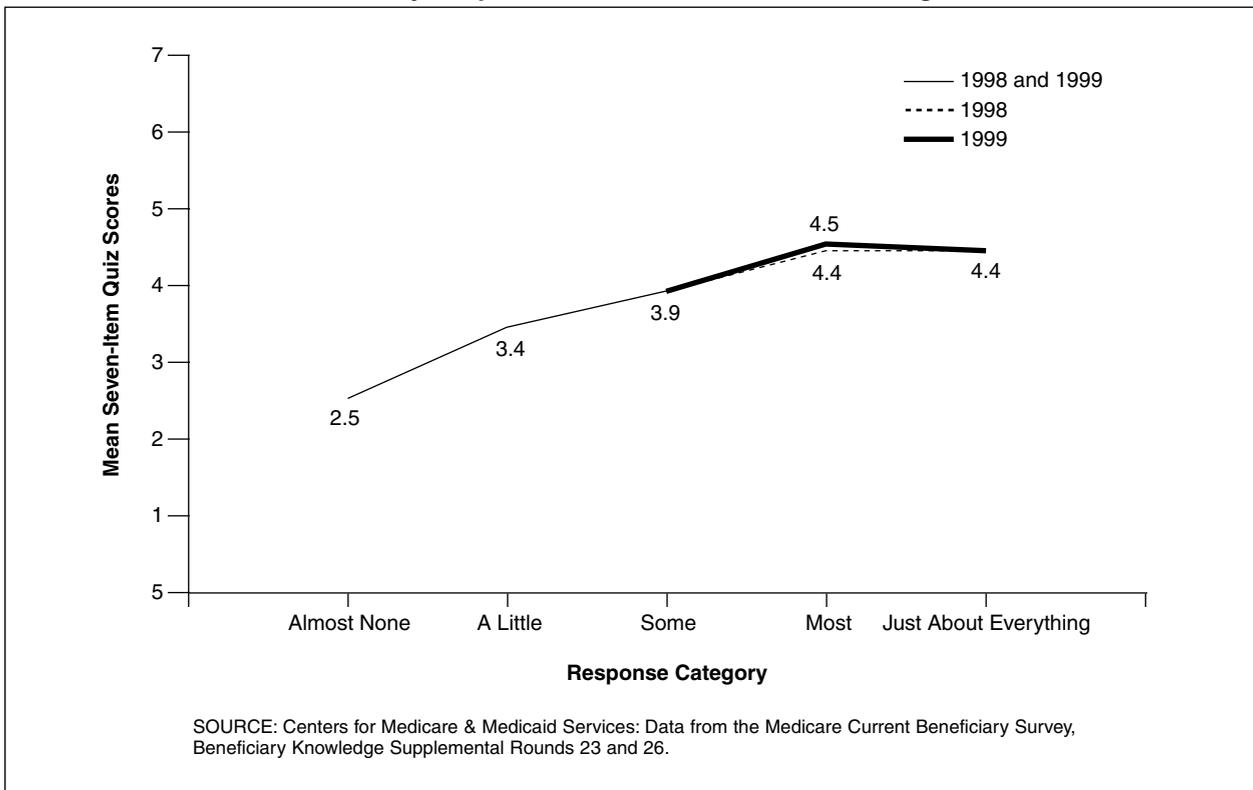### Relationship with Global Self-Perceived Knowledge

ANOVA results revealed that the quiz scores of sample members varied significantly, according to responses to the global self-perceived knowledge question in both 1998 ($F(4, 12972) = 348.68$, $p < 0.0001$) and 1999 ($F(4, 3909) = 119.86$, $p < 0.0001$). The mean seven-item quiz scores by global self-perceived knowledge appear in Figure 2. As shown in the figure, the means were identical in 1998 and 1999 for all levels except "most" and this category had only a 0.1 difference between the two years. For both years, there is a clear relationship between quiz scores and ratings on the global self-perceived knowledge question; respondents who perceived their knowledge to be higher demonstrated higher actual knowledge on the seven-item quiz. In both years, respondents with ratings of "most" or "just about everything" have similar quiz scores. It is possible that respondents cannot clearly distinguish between these two response categories.

### Group Comparisons

As shown in Table 6, seven-item quiz scores in 1998 differed significantly on all of the background variables, except for institutional utilization. The results for 1999 were significant for all variables, except institutional utilization, Part B utilization, and covered institutional charges. Examining the patterns of mean scores for

SOURCE: Centers for Medicare & Medicaid Services: Data from the Medicare Current Beneficiary Survey, Beneficiary Knowledge Supplemental Rounds 23 and 26.

both 1998 and 1999 indicated that the results are generally in the expected direction. Higher quiz scores were associated with sample members who were between the age of 65 and 75 or who had higher education, higher incomes, private supplemental insurance, or some charges. In addition, as hypothesized, respondents who were enrolled in managed care during the past year had higher seven-item quiz scores than those who were not enrolled during the past year.

## DISCUSSION

This study describes the development and psychometric evaluation of two potential knowledge measures, using data from the beneficiary knowledge and beneficiary needs supplemental rounds of the 1998 and 1999 MCBS. The first measure, the perceived knowledge index, includes five questions that ask beneficiaries to rate how much they know about particular topics related to Medicare. The other measure is a seven-item quiz that requires participants to respond to a series of true/false questions.

We evaluated the psychometric properties of the knowledge measures by calculating item- and scale-level descriptive statistics, assessing internal consistency reliability, and conducting construct validity analyses. Both the seven-item quiz and the perceived knowledge index exceeded the criterion for acceptable internal consistency reliability. The construct validity results indicated that both indices were significantly related to global self-perceived knowledge. Respondents who rated their knowledge higher on the global self-perceived knowledge question received higher scores on each index. In particular, there appeared to be a strong relationship

**Table 6**

**Subgroup Comparisons for Seven-Item Quiz Scores: 1998 and 1999**

| Variable | 1998 | 1999 |
|---|---|---|
| Education | $F_{(2,13059)} = 404.28$*** | $F_{(2,3904)} = 130.12$*** |
| Age Category | $t_{(11121)} = 13.22$*** | $t_{(3189)} = 8.33$*** |
| Income Category | $t_{(8539)} = -24.32$*** | $t_{(3075)} = -13.41$*** |
| Managed Care | $t_{(3834)} = -25.38$*** | $t_{(3918)} = -13.32$*** |
| Private Supplemental Insurance | $t_{(12140)} = -18.67$*** | $t_{(2040)} = -8.10$*** |
| Institutional Utilization[1] | $t_{(10662)} = 0.97$ | $t_{(3143)} = -0.92$ |
| Part B Utilization[1] | $t_{(1516)} = -7.08$*** | $t_{(3143)} = -1.68$ |
| Total Covered Charges[1] | $F_{(3,10660)} = 20.21$*** | $F_{(3,3141)} = 3.53$* |
| Covered Institutional Charges[1] | $F_{(3,10660)} = 3.88$** | $F_{(3,3141)} = 1.75$ |
| Allowed Part B Charges[1] | $F_{(3,10660)} = 21.33$*** | $F_{(3,3141)} = 2.86$* |

* $p < 0.05$.

** $p < 0.01$.

*** $p < 0.001$.

[1] Respondents enrolled in managed care during the past year were excluded from these analyses.

SOURCE: Centers for Medicare & Medicaid Services: Data from the Medicare Current Beneficiary Survey, Beneficiary Knowledge Supplemental Rounds 23 and 26.

between the PKI and the global self-perceived knowledge question, possibly because of the similarity in wording between the two.

In addition, both the PKI and the seven-item quiz performed well on the group validity comparisons. In general, groups previously shown to differ in knowledge had significantly different knowledge index scores. One of the most consistent findings across both years was a strong relationship between higher educational achievement and higher knowledge scores. Future research should explore possible implications of this relationship. If research shows that beneficiaries with lower educational achievement have difficulty understanding the Medicare educational materials, perhaps their knowledge could be improved by simplifying these materials. Another possibility is that those with higher education are more likely to read and retain the educational materials. The MCBS contains questions about reading and using the *Medicare & You* handbook which could be used for evaluating this hypothesis. Finally, it is possible that beneficiaries with higher education were able to deduce the correct answers from the wording of the items.

Based solely on the quantitative results, the PKI seems to have the best psychometric properties of the knowledge indices we examined. This index performed well in both the reliability and validity analyses, suggesting that it is the more precise measure of the construct it is measuring. However, other criteria, such as the content of the items, must be considered when selecting the most appropriate measure of knowledge. The PKI relies on beneficiaries to be the sole judge of their knowledge. Individuals' subjective ratings of their own knowledge may be influenced by factors other than knowledge, such as confidence in decisionmaking or satisfaction with information received. For example, beneficiaries who have someone else to make their health care decisions may feel they do not need to know any more information. Therefore, they could score themselves high on the PKI, regardless of what their actual knowledge level is. This possibility suggests that the PKI may not be the most accurate measure of actual knowledge.

A more objective measure of knowledge would require respondents to demonstrate their knowledge, such as by correctly answering the true/false questions included on the seven-item quiz. On the basis of

both content considerations and the psychometric analysis results, the seven-item quiz appears to be the more appropriate measure of beneficiary knowledge overall. The quiz had good variability in scores, reached an acceptable level of internal consistency reliability, and performed well on the validity analyses.

Some possible modifications could enhance the usability of the seven-item quiz. To obtain the most precise estimate of beneficiaries' knowledge possible, the quiz should contain questions that cover all difficulty levels. The grade-level analyses suggest that the items on the seven-item quiz cover all education levels, except for no formal education. Very few respondents report having no formal education; therefore, it may not be necessary to add questions specifically targeted at this group. However, the quiz contains only one item at the high school graduate level. Because more respondents in the MCBS report being a high school graduate than any other educational level, it may be helpful to add more items targeted to this group.

Furthermore, a limitation of the seven-item quiz is that four of its seven questions deal with managed care plans; this limitation restricts the range of knowledge the quiz can measure. Other questions could be added to improve the comprehensiveness of the quiz. For example, periodically, additional knowledge questions have been included in other rounds of the MCBS that could be used to expand the content coverage of the seven-item quiz. In addition, several new knowledge questions have been developed for the MCBS that address a variety of issues, including beneficiary rights and health plan decisionmaking (Uhrig et al., 2001).

Finally, future research should focus on developing additional versions of the knowledge quiz. Beneficiaries who have taken the quiz once may become more aware of the topics on the quiz and learn the answers to the questions. If the same set of questions is used every time, individuals may appear to have higher levels of knowledge, simply because they have been exposed to the questions before. The development of several equivalent forms of the knowledge quiz would enable researchers to administer different forms at each subsequent interview, thereby increasing the ability of the quiz to accurately measure change in knowledge over time.

## REFERENCES

Bann, C.: *Item Response Theory Analyses of the Demonstrated Knowledge Items From the 1999 Medicare Current Beneficiary Survey*. RTI. Research Triangle Park, NC. 2001.

Bann, C., and Berkman, N.: *Development and Psychometric Evaluation of Beneficiary Knowledge Indices from the Medicare Current Beneficiary Survey*. RTI. Research Triangle Park, NC. 2002.

Bann, C., Lissy, K.S., Keller, S., et al.: *Analysis of Medicare Beneficiary Baseline Knowledge Data from the Medicare Current Beneficiary Survey: Knowledge Index Technical Note*. RTI. Research Triangle Park, NC. 2000.

Cafferata, C.: Knowledge of Their Health Insurance Coverage by the Elderly. *Medical Care* 22(9):835-847, 1984.

Chapman, D.: A Survey of Non-Response Imputation Procedures. *American Statistical Association Proceedings, Section on Survey Research Methods*. American Statistical Association. Washington, DC. 1976.

Cronbach, L.J.: Coefficient Alpha and the Internal Structure of Tests. *Psychometrika* 16(3):297-334, 1951.

Gibbs, D.A., Sangl, J.A., and Burrus, B.B.: Consumer Perspectives on Information Needs for Health Plan Choice. *Health Care Financing Review* 18(1):55-74, 1996.

Guilford, J.P.: *Psychometric Methods*. McGraw-Hill. New York, NY. 1956.

Hibbard, J.H., Jewett, J.J., Englemann, S., and Tusler, M.: Can Medicare Beneficiaries Make Informed Choices? *Health Affairs* 17(6):181-193, 1998.

Howard, K.I., and Forehand, G.C.: A Method for Correcting Item-Total Correlations for the Effect of Relevant Item Inclusion. *Educational and Psychological Measurement* 22(4):731-735, 1962.

Hoyle, R.H., and Panter, A.T.: Writing About Structural Equation Models. In Hoyle, R.H. (ed.): *Structural Equation Modeling: Concepts, Issues, and Applications*. Sage Publications. Thousand Oaks, CA. 1995.

Hu, L., and Bentler, P.M.: Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. *Structural Equation Modeling* 6(1):1-55, 1999.

Jöreskog, K, and Sörbom, D.: *LISREL 8: User's Reference Guide*. Scientific Software International, Inc. Chicago, IL. 1996.

Lambert, Z.V.: Elderly Consumers' Knowledge Related to Medigap Protection Needs. *Journal of Consumer Affairs* 14(2):434-451, 1980.

Marquis, M.S.: Consumer's Knowledge About Their Health Insurance Coverage. *Health Care Financing Review* 5(1):65-80, Fall 1983.

McCall, N., Rice, T., and Sangl, J.: Consumer Knowledge of Medicare and Supplemental Health Insurance Benefits. *Health Services Research* 20(6): 633-657, 1986.

McCormack, L.A., Garfinkel, S.A., Hibbard, J.H., et al.: Health Insurance Knowledge Among Medicare Beneficiaries. *Health Services Research* 37(1):43-63, 2002.

McCormack, L.A., Anderson, W., Kuo, M. et al.: Measuring Beneficiary Knowledge in Two Randomized Experiments. *Health Care Financing Review* 23(1):47-62, Fall 2001.

Murray, L.A., and Shatto, A.E.: Beneficiary Knowledge of the Medicare Program. *Health Care Financing Review* 20(1):127-131, Fall 1998.

Nunnally, J.O.: *Psychometric Theory*. McGraw-Hill. New York, NY. 1978.

Satterthwaite, F.W.: An Approximate Distribution of Estimates of Variance Components. *Biometrics Bulletin* 2 (2):110-114, 1946.

Schumacker, R.E., and Lomax, R.G.: *A Beginner's Guide to Structural Equation Modeling*. Erlbaum. Mahwah, NJ. 1996.

Steel, R.G.D., and Torrie, J.H.: *Principles and Procedures of Statistics*, Second Edition. McGraw-Hill. New York, NY. 1980.

Uhrig, J.D., Squire, C., McCormack, L.A., et al.: *Questionnaire Development and Cognitive Testing Using Item Response Theory (IRT): Questionnaire Development Report*. RTI. Research Triangle Park, NC. 2001.

Ware, J.E., Snow, K.K., and Kosinski, M.: SF-36® *Health Survey: Manual and Interpretation Guide*. QualityMetric Incorporated. Lincoln, RI. 2000.