The Centers for Medicare & Medicaid Services' Center for Medicare (CM) strives to make information available to all. Nevertheless, portions of our files including charts, tables, and graphics may be difficult to read using assistive technology. Persons with disabilities experiencing problems accessing portions of any file should contact CM through e-mail at 508_Compliance@cms.hhs.gov.

MEDICARE HEALTH OUTCOMES SURVEY

FINAL REPORT

ON

IDENTIFYING ELDERLY HOS BENEFICIARIES AT RISK FOR MORTALITY USING THE UPDATED 2009 VES-HOS RISK SCORING

HEALTH SERVICES ADVISORY GROUP

MAY 2013

CMS CENTERS FOR MEDICARE & MEDICAID SERVICES CENTER FOR MEDICARE

HSAG HEALTH SERVICES ADVISORY GROUP

Table of Contents

List of Tables

Executive Summary

We extend and improve a previously validated Vulnerable Elders Survey-Medicare Health Outcomes Survey (VES-HOS) predictive model for identifying older adults at high risk of mortality using the HOS-2.0. In particular, we assess whether the approach is robust to the substitution of survey items. We extend the previously validated VES-HOS model (and the approach it is based on, the original VES-13 instrument) by incorporating an imputation approach that allows for the retention of a substantial proportion of the sample previously excluded because they were missing one or more items used by the algorithm; we verify the algorithm's performance by gender; we distinguish among three, rather than two, risk groups based on the risk score; and we characterize the members of vulnerable groups.

Background

One of the most widely used and cited instruments for identifying vulnerable elders is the 13-item VES-13 (Saliba, et al., 2001). Ng, Elliott, Scholle, Ahmed, Collins, and Bierman (2012) adapted the VES-13 by matching analogous VES-13 items to the 2005 HOS 1.0 instrument to identify Medicare Advantage Organization (MAO) beneficiaries aged 65 years and older at risk of dying within two years. However, in 2006 the HOS 1.0 instrument was updated to the HOS 2.0 with new questions and response categories that could affect the utility of the adapted VES-13 developed by Ng et al. (2012).

The VES-HOS Risk Score has four components: age (1 item), self-rated health (1 item), physical activities (4 items) and physical condition (5 items). The content and wording of age and self-rated health were virtually unchanged from the 2005 HOS 1.0, but the content and/or wording of many of the physical activities and physical condition items changed substantially. Of particular concern was that there were no corresponding questions in the 2009 HOS 2.0 for three of the four physical activity items used from the 2005 HOS 1.0. One (of five) physical condition item was dropped since 2005, and there were slight changes to question wording and/or response categories. Because of these changes, the 2005 version of the VES-HOS could no longer be employed, and it was unclear whether a modified VES-HOS scoring could be reliably applied to the most current (2009) HOS baseline data to identify vulnerable seniors.

We report here on our efforts to extend and improve the previously validated VES-HOS predictive model. In addition, we considered several potential innovations to the VES-HOS algorithm. The original VES-13 and the 2005 VES-HOS each excluded beneficiaries who were otherwise eligible to be assigned a risk score, except they were missing one or more items used in the scoring. This exclusion is particularly concerning because high rates of survey item non-response among seniors may be related to their vulnerability (De Leeuw, Joop, & Huisman, 2003; Elliott, Edwards, Angeles, Hambarsoomians, & Hays, 2005). Accordingly, we developed and validated an imputation approach that allows these cases to be assigned a risk score.

Gender was not explicitly incorporated in either the original or 2005 VES-HOS scoring. In the development of VES-13 scoring, item response theory was used to identify survey items from the Medicare Current Beneficiary Survey (MCBS) that demonstrated consistent predictive ability (of functional decline and death) across three population subgroups: men $\geq$65 years, women $\geq$65 years, and persons $\geq$85 years, but gender was not directly incorporated into the scoring (Saliba, et

al., 2001). Min, Elliott, Wenger, and Saliba (2006) used data from a health care plan to assess whether the VES-13 predicted both mortality and functional decline in a sample of seniors aged 75 and older for a one-year interval, and found that it did so. They also explored whether the inclusion of gender as a "main effect" in the original VES-13 scoring improved its ability to predict health decline, and found that it did not. Min et al. (2006) did not explore gender as a potential moderator of the VES-13 score, i.e., the possibility that the relationship between scores and outcomes differed by gender in slope rather than intercept. No one has yet examined the question about whether including gender improved the predictive ability of the VES-HOS algorithm.

Finally, we identified and described the socio-demographic characteristics and patterns of sources of points of three risk groups based on risk scores. In the original VES-13 validation study, a "cutoff" score of ≥3 based on sensitivity analyses distinguished high risk vulnerable older adults with 4.2 times the risk of two-year mortality or functional decline compared to low risk vulnerable adults (Saliba, et al., 2001). We explored whether three rather than two risk groups with distinctive socio-demographic and health profiles can be identified to better differentiate the risk for mortality. As in the 2005 VES-HOS study, the 2009 VES-HOS algorithm could not determine risk for functional decline.

Instruments and Data Source

The HOS was first fielded nationally in 1998, and is the first patient-based outcomes measures in Medicare managed care. The HOS is a longitudinal survey that assesses the physical and mental functioning of beneficiaries.

We used data from beneficiaries who responded to the 2009 HOS 2.0, which obtained a sample from all MAOs with a minimum enrollment of 500 members. MAOs with Medicare contracts in effect on or before January 1, 2008, including local and regional preferred provider organizations (PPOs) and continuing cost contracts, participated in the HOS Baseline Survey in 2009. MAOs composed exclusively of special needs plan (SNP) benefit packages, regardless of institutionalized, chronically ill or dual eligible enrollment, were also included in the requirement. Private fee-for-service (PFFS) plans could voluntarily report HOS in 2009 (NCQA, 2009).

A total of 238,687 respondents (from 424 MAO contracts) were analyzed. For the purposes of this report, all analyses included observations from seniors, aged 65 and older, who had completed at least 80% of the 2009 Cohort 12 Baseline HOS items (or 51.7% of eligible seniors from the 2009-2011 Cohort 12 Merged Baseline and Follow Up file).

Methods

We computed a risk score based on age, self-rated health, limitations on physical functioning (getting in/out of chairs, accomplished less than liked to because of physical problems, difficulty climbing stairs), and disability (number of days physical health not so good, feeling depressed/sad, accomplished less due to emotional problems, difficulty walking, limitation in moderate activities, physical health limiting kind of work/activities, difficulty bathing) for HOS baseline beneficiaries aged 65 and older who completed at least 80% of survey questions. We

used a modified predictive model that substituted for several items used in two previously validated versions of the risk score.

We compared risk score patterns and distributions to those obtained using a previously validated version. We estimated multivariate models to predict two-year mortality rates from risk score, age, sex, race/ethnicity, marital status, education, annual household income, and Medicaid indicator; and compared the risk score and two-year mortality rate with that reported previously by Ng et al. (2012) using the HOS 1.0 instrument.

We compared the two-year mortality rate of the unimputed cases to those with any missing items (combined) and by number of missing items. We used median imputation and two alternate (most positive response and least positive response) forms of imputation for missing item scores. We compared the distribution of risk scores and the relationship between risk score and two-year mortality for the imputed sample under each imputation approach against the unimputed sample to identify the best imputation method.

Using the full sample (incorporating what was determined to be the best imputation approach), we compared the frequency of 2009 VES-HOS Risk Scores by gender, the odds of death for each one-unit increase in the risk score by gender, and the mortality rates by gender for each risk score. Gender specific mortality rates were stratified by age groups (e.g., 65-74, 75-84, and 85 or more years) in order to examine whether age further differentiated the risk for mortality by gender as shown in the 2009 VES-HOS Risk Score algorithm.

Also using the full sample, we defined three risk groups: low risk (0-2 points), moderate risk (3-6 points), and high risk (7-10 points). We examined the distribution of socio-demographic characteristics within each risk group. We used a series of logistic regressions predicting each row indicator from a dummy for moderate risk and a dummy for high risk (low risk was excluded) to identify significant group differences. We also described the percentage of beneficiaries in each risk group who obtained their points from each proxy HOS 2.0 item used in the algorithm.

Key Findings

For each one-point increase in the 0-10 risk score, 2009 HOS baseline respondents had 37% higher odds of dying within two years ($p<0.0001$). One in ten (10.6%) of those scoring ≥ 3 died within two years, compared to 2.4% of those scoring <3, resulting in a relative risk of death of 4.4, similar to the original VES-13 sample (where scores ≥ 3 indicated 4.2 times the risk of death or decline, relative to scores <3). There were 15.5% of those scoring ≥ 7 who died within two years, resulting in a relative risk of death (relative to scores <3) of 6.5 (not shown in tables).

The two-year mortality rate for those with missing items is 9.5% compared to 7.1% for those with no missing items; however, this is not a dose-response relationship, in that mortality is higher for those with any item missingness than with none, but it does not increase with the number of missing items among those with any missing items. We find that the mortality risk for unimputed cases is best predicted using median imputation for missing items.

We find that on average, women score 0.5 to 2 points higher than men on risk scores (varying slightly with the range of risk scores), largely because of the greater average age of women relative to men among MAO beneficiaries aged 65 and older (and among US seniors more generally). Nevertheless, we find that the same risk score predicts the same risk of death and decline for men and women.

We find that 38% of persons aged 65 and older are at low risk (0-2 points), 29% are moderate risk (3-6 points), and 33% are high risk (7-10 points). Compared to low risk beneficiaries, moderate and (especially) high risk beneficiaries are successively characterized as follows:

- Older
- More often female
- More often Black and other/unknown race/ethnicity
- Less often married and more often widowed
- Less educational attainment
- Lower household income

Age is the primary source of points in the low-risk group and plays a role in distinguishing the high-risk group from the other two. Self-rated health plays a role in distinguishing the high-risk group from the other two groups, though it is not a primary source of points for any group. Physical activity distinguishes each of the three groups. Physical condition distinguishes the moderate and high-risk groups from the low-risk group, is the largest source of points for both moderate and high-risk groups, and accounts for two-thirds of all points in the moderate-risk group.

Implications

We were able to update the Ng et al. (2012) 2005 VES-HOS 1.0 algorithm for use with the most recent (2009) HOS baseline data, suggesting that future changes to the HOS can be incorporated into the VES-HOS algorithm.

We expanded the utility of the 2009 VES-HOS Risk Score so that it is now applicable to the 16% of cases that would have otherwise been lost to missing data, which includes a disproportionate number of vulnerable beneficiaries. We also developed a sub-classification of a very high-risk group than was previously available, and we validated this sub-classification. The over-representation of women, racial/ethnic minorities, and those with low socio-economic status (SES) in the highest-risk group emphasizes the need for continued attention to risk factors for preventable death for these beneficiaries. These improvements and innovations increase the ability of the algorithm to allow the Centers for Medicare & Medicaid Services (CMS) and MAOs to identify the most vulnerable beneficiaries and to address their health care needs.

Health care organizations can use the updated 2009 VES-HOS with the HOS baseline data to assess the degree of mortality risk within its older patient population, as well as to target specific members for quality improvement intervention. MAOs can encourage health practitioners to use the updated 2009 VES-HOS as a real-time screening tool for mortality risk among their elderly beneficiaries. Additionally, MAOs can retrospectively apply the 2009 VES-HOS algorithm to

their HOS baseline data and analyze the standard of care for those beneficiaries with greater risk scores over the preceding two years.

Limitations

The HOS involves a 2-year follow up; therefore, we do not know the extent to which the 2009 VES-HOS 2.0 algorithm predicts mortality within a shorter or longer time period, nor do we know the extent to which it predicts elders at risk of functional decline. However, prior work using the original VES-13 shows that the VES-13 is predictive of both mortality and functional decline within shorter and longer time intervals.

The HOS is restricted to the MAO population; therefore, we cannot be certain if the results apply to a more general beneficiary population without further validation of the VES-HOS in a different population. However, we note that the original VES-13 was validated using a 1993 national survey of community-dwelling beneficiaries, prior to the advent of Medicare Advantage.

MAOs do not receive their HOS baseline data until the two-year follow up period is scored. Therefore, the 2009 VES-HOS algorithm cannot be applied to baseline data to prospectively identify at-risk beneficiaries for appropriate health care interventions.

Future Work

Since we were able to validate the VES-HOS Risk Score using the 2009 HOS baseline data, future work can use these risk scores to address substantive questions. Examples of such analyses include examining the distribution of high and moderate risk beneficiaries across contracts and developing feedback to contracts about these distributions. In addition, we could examine how quality of care, as measured by the HOS, differs by risk status. This could be done both overall and by contract. Analyses of the latter type could determine the extent to which MAO performance for less vulnerable beneficiaries is predictive of their performance for more vulnerable beneficiaries. This would provide evidence of whether some MAOs appear to offer specialized care and have specific areas of organizational strengths or weaknesses that might inform quality improvement efforts.

Chapter 1: Introduction

Purpose

The primary purpose of this report is to provide the results of an effort to update and validate the modified 2005 Vulnerable Elders Survey-Medicare Health Outcomes Survey (VES-HOS) 1.0 algorithm to match the HOS 2.0 instrument using the baseline responses and two-year follow up death status from the 2009-2011 HOS Cohort 12 Merged Baseline and Follow Up file. We also investigate three potential extensions of the VES-HOS algorithm. First, we investigate potential imputation approaches that would assign risk scores to seniors who previously did not receive a score because of item non-response. Second, we evaluate the algorithm's performance by gender. Third, we identify three risk groups, rather than two as done in the original VES-13 and 2005 VES-HOS. We also describe the socio-demographic characteristics of these three risk groups and describe the age and health characteristics that largely determine assignment to each of the three risk groups.

Background

In order to improve performance, Medicare Advantage Organizations (MAOs) are seeking ways to identify at-risk beneficiaries who would benefit from outreach and intervention. The HOS contains important information to help guide decision-making in identifying such approaches likely to be effective. We use baseline data from the HOS 2009-2011 Cohort 12 Merged Baseline and Follow Up file to update a predictive model to identify the at-risk population. With this model, we are able to provide MAOs with a simple, easily-implemented screening tool to identify vulnerable MAO beneficiaries.

One of the most widely used and cited instruments for identifying vulnerable elders is the 13-item VES-13 (Saliba, et al., 2001). This instrument is able to predict hospitalizations, death, functional decline, and resource use among older adults (Pacula, Boult, Reed, & Aliberti, 1997), and is able to identify seniors at high risk of death or decline. However, the data elements needed to calculate the VES-13 are not routinely available for MAO enrollees. Ng et al. (2012) adapted the VES-13 by matching items from the 2005 HOS 1.0 to analogous items in the VES-13 to identify MAO beneficiaries aged 65 and older at risk of dying within two years. Each year, the HOS measures a range of beneficiary characteristics, including socio-demographics, health-related quality of life (HRQOL), limitations in activities of daily living (ADLs), and information about depression risk. However, the content of the HOS instrument has changed over time. Because of these changes, the 2005 version of the VES-HOS could no longer be employed with the 2009 HOS 2.0 instrument, and it was unclear whether a modified VES-HOS scoring could be reliably applied to the most current (2012) HOS baseline data to identify vulnerable seniors.

We investigate whether the 2009 HOS 2.0 instrument can be used to accurately identify at-risk elders. Because the 2012 HOS instrument is the same as the 2009 instrument (with the exception of changes to the height and weight items, neither of which are used in the VES-HOS scoring), our results have implications as to whether MAOs can also use 2012 HOS data to identify at-risk members.

This report explores three potential innovations to the VES-HOS Risk Score approach. First, we develop and validate a principled, easy-to-implement imputation approach that allows for the

inclusion of seniors who previously would have been excluded because they did not provide information on all the items used in the scoring. This benefits the HOS in two ways. First, it improves sampling efficiency by allowing more beneficiaries to be included in the analytic sample and more vulnerable beneficiaries to be identified. Second, retaining these cases improves sample representativeness. Because high rates of survey item non-response among seniors may be related to poorer health status, excluding seniors with missing information potentially biases the observed risk scores and disproportionately fails to detect vulnerable seniors (De Leeuw, Joop, & Huisman, 2003) (Elliott, Edwards, Angeles, Hambarsoomians, & Hays, 2005). We investigate and find evidence to suggest that seniors with missing VES-HOS items have higher mortality rates than those with no missing information, confirming the concern; but establish that the imputation approach addresses this concern. We conclude this analysis by recommending that users employ median imputation of missing values to retain seniors with any missing VES-HOS items.

Second, we verify the algorithm's performance by gender. Gender was not explicitly incorporated in the original VES or the 2005 VES-HOS scoring. In the original development of VES-13 scoring, item response theory was used to identify survey items from the Medicare Current Beneficiary Survey (MCBS) that demonstrated consistent predictive ability (of functional decline and death) across three subgroups: men $\geq$65 years, women $\geq$65 years, and persons $\geq$85 years, but gender was not directly incorporated into the scoring (Saliba, et al., 2001). Min et al. (2006) used data from a health care plan to assess whether the VES-13 predicted mortality and functional decline over a one-year interval in a sample of plan members aged 75 and over. As part of this assessment, they also explored whether the inclusion of gender as a "main effect" in the original VES-13 scoring improved its ability to predict health decline, but found no significant effect. Min et al. (2006) did not explore gender as a potential moderator of the VES-13 score (i.e., the possibility that the relationship between scores and outcomes differed by gender in slope rather than intercept). The question regarding whether gender should be incorporated in the VES-HOS scoring had not been assessed prior to our report.

Third, we extend the number of distinct risk groups from two to three. In the original VES-13 validation study, a "cutoff" score of $\geq$3 based on sensitivity analyses distinguished high risk vulnerable older adults with 4.2 times the risk of two-year death or decline compared to low risk vulnerable adults (Saliba, et al., 2001). We further distinguish between vulnerable seniors who have moderate risk from those with high risk of two-year mortality. We are able to do so because we have a larger base sample and analytic sample (due to imputation) than used in the development and validation of either the original VES-13 or the 2005 VES-HOS.

Our report also describes the socio-demographic characteristics of the three risk groups (low, moderate, and high) identified in the 2009 HOS baseline data using the 2009 VES-HOS Risk Scoring. We also investigate the key characteristics such as age, self-rated health, physical activities, and physical condition that characterize each of the three groups.

Chapter 2: Methods

Survey Instrument

Data from the HOS 2009-2011 Cohort 12 Merged Baseline and Follow Up were used for these analyses. The HOS is a longitudinal survey that assesses the physical and mental functioning of the aged and disabled beneficiaries in MAOs over a two-year period (baseline and follow up surveys). Survey vendors are certified each year by the National Committee for Quality Assurance (NCQA) and follow the NCQA guidelines (NCQA, July 8, 2009). A description of the HOS 2.0 instrument appears below and copies may be accessed from the www.HOSonline.org website.

The HOS survey was first implemented in 1998 by the Centers for Medicare & Medicaid Services (CMS) to measure a health plan's ability to maintain or improve the physical and mental health of its beneficiaries over time (Haffer & Bowen, 2004). The HOS is administered annually to a random sample of individuals drawn from all plan benefit packages of each participating MAO. Each spring a baseline survey is administered to a new cohort of Medicare beneficiaries. Each cohort of beneficiaries is resurveyed in two years.

MAO contracts are required to have a minimum of 500 beneficiaries to participate in the HOS baseline. For MAOs with 1,200 or more beneficiaries, a random sample of 1,200 is drawn, and for contracts that have at least 500 beneficiaries but less than 1,200, all beneficiaries are sampled. MAOs with a minimum enrollment of 500 beneficiaries, including local and regional preferred provider organizations (PPOs), and continuing cost contracts that held §1876 risk or cost contracts with Medicare contracts in effect on or before January 1, 2008, were required by CMS to participate in the HOS Cohort 12 Baseline survey in 2009. MAOs composed exclusively of Special Needs Plans (SNPs) benefit packages, regardless of institutionalized, chronically ill or dual eligible enrollment, are also included in this requirement. Some Private Fee-for-Service (PFFS) contracts voluntarily reported the HOS in 2009 (NCQA, HEDIS Specifications for the Medicare Health Outcomes Survey, 2009).

In 2006, CMS implemented the Medicare HOS 2.0 for MAOs (NCQA, 2009). The HOS 2.0 evaluates the HRQOL of MA beneficiaries by measuring their physical and mental health status using the Veterans RAND 12-Item Health Survey (VR-12) (Iqbal et al., 2007), instead of the 36-item health survey used in the HOS 1.0. The HOS 2.0 also contains questions about socio-demographics, ADLs, chronic medical conditions, depression risk, and height and weight used for calculation of Body Mass Index (BMI). Four HEDIS® Effectiveness of Care measures are included to evaluate management of urinary incontinence, physical activity, osteoporosis testing and fall risk management.

The HOS is a patient-reported survey with mail (two survey mailings) and telephone components. Survey vendors attempt telephone follow up with at least six attempts in those instances when beneficiaries fail to respond after the second mail survey. The survey has English, Spanish, and Chinese language versions available. The present analyses use baseline surveys and death status at two-year follow up.

Data Source (Table 1)

The HOS data we employed included a wide range of variables such as those related to socio-demographics, HRQOL measures, limitations in ADLs, and information about depression risk. The study sample was drawn from the HOS 2009-2011 Cohort 12 Merged Baseline and Follow Up file, which included one baseline record per beneficiary (n=487,861) from 424 MAO contracts. Of those eligible to participate in the HOS (n=461,699), the study was restricted to seniors aged 65 and older who had completed 80 percent or more of the HOS (n=238,687). Death status at two-year follow up was obtained from the CMS database by RTI International and merged to the HOS by the Medicare Health Insurance Claim (HIC) Number. The restrictions to the study sample were developed to match the sample defined by Ng et al. (2012) and, thus, allow comparison of results between the two studies. Table 1 shows that 92.5% (n=220,818) of seniors were still alive at the two-year follow up and 7.5% (n=17,869) had died. As expected, non-survivors were generally older than survivors, and had lower socio-economic status (SES) as measured by education, annual household income, and Medicaid status compared to survivors. Females made up a disproportionate percentage of survivors compared to males (58.0% versus 42.0%), but this gap narrowed considerably among non-survivors (51.7% versus 48.3%). An examination of marital status shows that 56.6% of survivors were married and 27.1% were widowed. Among non-survivors, marital status declined to 46.0% and widowhood increased to 38.8%. There were no noticeable differences in the racial breakdown among survivors and non-survivors.

Variable Descriptions (Table 2)

We examined the extent to which the content of the 2009 HOS 2.0 has changed since the 2005 HOS 1.0 and the survey items used in the original VES-13. Table 2 describes the predictor variables used for the original VES-13, the 2005 VES-HOS 1.0 algorithms and the variables proposed to be used for the 2009 VES-HOS 2.0 algorithm. Where possible, we identified the questions and responses identified from the 2009 HOS 2.0 that are most similar in wording and content to those used in the 2005 HOS 1.0. We identified possible substitutes that appeared to capture the content of items eliminated since the 2005 HOS 1.0. The 2009 HOS 2.0 includes questions on age, self-rated health status, and two function-based items (ADL disability in walking and bathing) that were similar to those used in the original VES-13 and 2005 VES-HOS algorithm modified for use with the 2005 HOS 1.0. The 2009 HOS 2.0 also includes two function-based items: limitation in moderate activities (pushing a vacuum cleaner, moving a table, bowling, playing golf) and limitation in work/activities due to physical health. Finally, the 2009 HOS 2.0 includes two measures capturing mental functioning (presence of depressed mood in past year and limitation in work/activities due to emotional health) that were used in the 2005 HOS 1.0 as substitutes for items found in the original VES-13 scoring.

There were items from the 2009 HOS 2.0 that we considered as substitutes for four previously used items. The 2009 HOS 2.0 includes one function-based item that addresses similar content to that used in 2005 VES-HOS 1.0 and original VES-13: limitation in climbing several flights of stairs (possible substitute for limitation in walking several blocks). As a possible substitute for limitation in lifting groceries or vigorous activities, we explored accomplished less due to physical problems; for lifting groceries item, we explored both physical or mental health limits usual activities and number of days in past 30 days physical health not good. For bending, kneeling, or stooping, we explored getting in/out of chairs.

In the original VES-13 validation study, the 13 predictor variables selected for the instrument collectively predicted high risk of death or functional decline over two years (decline defined as [1] change from no Instrumental Activity of Daily Living (IADL) or ADL disability to any IADL or ADL disability, [2] an increase of two or more in the total count of IADL or ADL disabilities, or [3] new admission to a nursing home). However, the 2009 HOS 2.0 did not specifically assess IADL disability or nursing home admission. Because of this, and given that the original VES-13 predictors performed comparably well in assessing death or decline, only the death outcome was employed in the 2005 HOS 1.0. We employed a similar approach to develop and validate our prediction model of mortality based on the 2009 HOS 2.0. This approach is consistent with prior techniques for developing database-derived proxies of a survey instrument, which have employed an outcome identical to, not just similar to, an outcome used in the instrument's original validation study (Coleman, Wagner, Grothaus, Hecht, Savarino, & Buchner, 1998; Vojta, TenHave, Amaya, Lavizzo-Mourey, & Asch, 2001).

Validation Approach (Table 3)

We conducted three sets of analyses to validate the proposed updated algorithm. First, we compared the distribution of the variables in the variable list with the variable distributions reported by Ng et al. (2012) in the validation using 2005 HOS 1.0 variable lists and data.

Second, we compared the distribution of the 2009 VES-HOS Risk Score using the proposed variable list and scoring (as described in Tables 2 and 3) with the risk score distribution reported by Ng et al. (2012). In the proposed scoring, age of 75-84 is awarded one point, while seniors aged 85 and older receive three points. One point is awarded for reporting fair or poor health. One point is awarded for each of the following: have difficulty or unable to get in/out of chairs, accomplishing less (some, most, or all of the time) due to physical problems, limited a lot in climbing several stairs, and limited a lot in moderate activities. Four points are awarded for one or more responses to: number of days physical health not so good (14-30), feeling depressed/sad (yes), accomplished less (some, most, or all of the time) due to emotional problems, difficulty walking (any or unable to do), physical health limiting kind of work/activities (some, most, or all of the time), and difficulty bathing (any or unable to do). These two analyses provided an assessment of the equivalence of the distribution of the proposed variables and scoring with that described by Ng et al. (2012).

Third, we compared the relationship between 2009 VES-HOS Risk Scores and 2011 follow up mortality with the relationship between the 2005 VES-HOS Risk Scores and 2007 follow up mortality.

Extending the Algorithm to Incorporate Missing Cases and Evaluating Treatment of Gender-Specific Patterns

Neither the original VES-13 nor the 2005 VES-HOS Risk Score methods included survey respondents with missing items as were incorporated in the updated 2009 VES-HOS scoring. With samples of older adults, exclusion based on item non-response can be problematic since the oldest and most vulnerable seniors are among those most likely to have item non-response. We first examined whether there is evidence that cases with missing values differ in mortality from cases without missing values by comparing the two-year mortality rate for the 199,992

beneficiaries with a risk score under standard approaches, and the 38,695 cases with any 2009 VES-HOS item non-response.

Next, we compared three different imputation approaches to determine which best predicted mortality. We considered three possible models for missing item imputation: 1) imputing the most positive, 2) imputing the median, and 3) imputing the least positive value for missing items. We performed a series of analyses to identify the best imputation method. We compared the distribution of risk scores and of the relationship between risk score and two-year mortality for the imputed sample under each imputation approach. We also compared the odds ratios associated with a one-unit increase in the risk score for the unimputed sample and the imputed sample under each of the three strategies. Finally, we considered the calibration of the death rates at each risk score for the unimputed sample and the imputed sample using each of the three strategies.

We also examined the algorithm's performance by gender. We compared men's and women's two-year mortality rates and average ages to determine whether there are likely gender differences in risk scores. We compared the frequency of 2009 VES-HOS Risk Scores by gender, the odds of death for each one-unit increase in the risk score by gender, and the mortality rates by gender for each risk score. Gender specific mortality rates were stratified by age groups (e.g., 65-74, 75-84, and 85 or older) to assess whether age further differentiated the risk for mortality by gender as shown in the 2009 VES-HOS Risk Score algorithm.

Defining and Describing the Characteristics of Each of Three VES-HOS Risk Groups

In prior years, risk scores have been used to distinguish low-risk (<3 points) and high-risk (≥3 points) seniors. In these analyses we considered the viability of further distinguishing those at high risk (≥ 7 points) from those at moderate risk (3-6 points). We described the characteristics of the three risk groups. These analyses were based on the version of the 2009 HOS baseline data with imputed cases and involved a series of logistic regressions predicting each row indicator from a dummy for moderate risk and a dummy for high risk (low risk was excluded). To gain further insight into the characteristics of each of the three risk groups, we describe the percentage of beneficiaries in each risk group by the source of VES-HOS points.

Chapter 3: Results of Analyses to Update and Validate the VES-HOS Algorithm

This chapter describes the analyses used to support the development of the proposed VES-HOS scoring algorithm (i.e., comparison of the distribution of variables used in the 2005 VES-HOS Risk Score algorithm with the distribution of proposed variables for the 2009 VES-HOS Risk Score algorithm) and its validation (i.e., comparison of the distribution of VES-HOS Risk Scores in 2005 and 2009 and an examination of the relationship between two-year mortality and VES-HOS Risk Scores).

Findings

The results for the comparison of the distribution of the proposed 2009 VES-HOS scoring variables with the variables used in the 2005 VES-HOS scoring are shown in Table 4. In cases where we substituted a new survey item for a 2005 VES-HOS item, we used the comparison of distributions to select cut-off values for the new item. Table 5 compares the distribution of individual 2005 and 2009 VES-HOS Risk Scores. We found that the distributions of individual scores were remarkably similar. For further validation, we compared the relationship between the linear VES-HOS score and two-year mortality (Table 6) and assessed the pattern of the relationship between individual risk score point values and mortality in both years (Table 7). These relationships were quite similar for the 2005 and 2009 versions, indicating that the VES-HOS modification was successful and quite robust to the substitution of items across survey forms.

Results of Comparison of Distribution of VES-HOS Risk Score Variables (Table 4)

To assess proposed variable substitutions and cut-offs to use in the algorithm, we also compared the distribution of components of the 2005 and 2009 VES-HOS Risk Scores. Table 4 compares the values for HOS items that are the same in 2005 and 2009 and the values for proposed substitutes for HOS variables that were dropped from the HOS after 2005. We examined the distribution of values for variables proposed as 2009 substitutions for variables dropped after 2005 from the HOS to identify changes in items or scoring that could improve the correspondence between the 2005 and 2009 VES-HOS Risk Score distributions.

In general, the correspondence is high for the items with identical or very similar content and wording in 2005 HOS 1.0 and 2009 HOS 2.0 instruments (age, self-rated health, depressed mood, and the three functional items of moderate activity, walking, and bathing). The correspondence between the proposed substitute items (using proposed cut offs) for items dropped after 2005 is also generally high. For limitations in bending, kneeling, or stooping (24.6%), we substituted have difficulty or unable to get in/out of chairs (22.4%) and for limitation in lifting groceries (52.7%) we substituted accomplishing less (some, most, or all of time) due to physical problems (44.8%). For limitations in walking several blocks (55.7%), we substituted limited a lot in climbing several stairs (27.9%), which has the lowest correspondence to the dropped variable of any of the proposed substitutes. For limitations lifting groceries (16.2%), we substituted poor health for at least 14 of the last 30 days (19.9%). The substitutes for the items involving reduced time on work and activities due to emotional and physical health (22.4% and 37.9%, respectively) were items regarding limits in kind of work or other activities

that can be done some or more of the time due to emotional or physical health (22.4% and 44.4%, respectively).

Results of Comparison of Distribution of VES-HOS Risk Scores in 2005 and 2009 (Table 5)

We computed a VES-HOS Risk Score for every beneficiary in the 2009 HOS baseline eligible to receive a risk score (i.e., with no item non-response) using the variable list and scoring described in Tables 1 and 2. Table 5 shows the distribution of 2005 HOS sample in column 1 as presented in Ng et al. (2012) and the corresponding distribution using the 2009 HOS sample in column 2.

The 2009 distribution is similar to the 2005 distribution with 61.1% of the 2005 HOS sample receiving a risk score of 3 or higher compared to 58.5% of the 2009 HOS sample. This result provides initial evidence that the calibration of the proposed VES-HOS variable list and risk score algorithm for use with the 2009 HOS 2.0 is similar to that for the 2005 VES-HOS 1.0 developed by Ng et al. (2012).

Results of Examination of Relationship between Two-Year Mortality and VES-HOS Risk scores (Tables 6-7)

Next, we assessed how well the proposed scoring system and candidate items predicted two-year mortality in 2009 (1=died, else=0) using all available cases, and compared these results to the 2005 results reported by Ng et al. (2012). Since individuals are clustered within MAO contracts, mixed-effect (or hierarchical) multivariate logistic regression (with random effects for contracts) was used to examine the dependent variable of death over two years as a function of the 2009 VES-HOS Risk Score (range 0-10).

Table 6 shows that the risk score significantly predicted death within two years in 2009, as it did in 2005. Specifically, for each one-point increase in the 0-10 risk score, 2009 respondents had 37% higher odds of death ($p<0.0001$), compared to 36% higher odds of death per point observed in 2005.

To investigate whether the single 0-10 risk score and mortality is a linear relationship, we calculated the mortality rate for persons with each risk score. As shown in Table 7, as was the case in 2005, this relationship was not linear (though it was monotonically increasing) in 2009.

Since the pattern of the relationship between mortality and the risk scores is similar to that reported by Ng et al. (2012), we conclude that the proposed 2009 VES-HOS Risk Score algorithm is validated, and we recommend its use for identifying vulnerable seniors.

Chapter 4: Results of Analyses Related to Extending the VES-HOS Algorithm to Account for Missing Data and Investigating Gender-Specific Patterns

In this chapter, we explore issues related to imputing risk scores for seniors with any VES-HOS item non-response and who completed at least 80% of the HOS survey. We also investigate whether gender should be directly incorporated in the 2009 VES-HOS scoring.

Findings

We describe a key extension of the 2009 VES-HOS algorithm: a principled approach to incorporating cases with any missing 2009 VES-HOS measures into the sample (Tables 8-11) among those who completed at least 80% of the HOS survey. On the basis of the analyses, we recommend that median imputation be incorporated into the 2009 VES-HOS algorithm (and a similar approach may be useful for the parent VES-13). We also describe our investigations into whether the VES-HOS algorithm would benefit from incorporating gender (Tables 12-17). We conclude that the association between risk score and 2-year mortality differs slightly by gender, but that incorporating gender does not sufficiently improve the predictive ability of the VES-HOS algorithm to merit its inclusion.

Results of Assessment of Potential Bias from Excluding Missing Cases (Table 8)

In the absence of imputation, 16% (n=38,695) of beneficiaries who were eligible for the analytic sample (i.e., aged 65 and older and having completed 80% or more of the survey) were dropped from the sample because they were missing at least one item used in the 2009 VES-HOS scoring. To assess the potential bias that results from their exclusion from the sample, we compared the two-year mortality for seniors with missing items and those assigned a risk score without imputation. The two-year mortality rate for those with missing items was 9.5% compared to 7.1% for those with no missing items, suggesting that the exclusion of these seniors positively biases the observed health of the sample and omits a disproportionately vulnerable subset of seniors, lessening the instrument's value as a screening tool.

Results of Assessment of Three Imputation Models (Table 9-11)

We found that beneficiaries with any missing VES items have higher mortality rates than those with none among those who completed 80% or more of the overall survey. However, there is not a dose-response relationship (see Appendix Table A1), in that mortality is higher for those with any item missingness than with none, but it does not increase with the number of missing items. In addition, we find that the mortality risk for unimputed cases is best predicted using median imputation for missing items. We considered three different ways to impute scores: median imputation and two alternate forms of imputation for missing item scores (most positive response option and least positive response option).

Table 9 shows the distribution of risk scores for the 199,992 unimputed beneficiaries and compares this distribution with the distribution of scores for those for whom imputation was necessary when imputed under each of the three imputation strategies: median, most positive response, and least positive response. Under each strategy, the health distribution of beneficiaries

with any missing values tends to be worse. This is consistent with the higher two-year mortality rate among those with any missing values relative to those without missing values.

Table 10 shows the odds ratio associated with a one-unit increase in the 2009 VES-HOS Risk Score for the unimputed sample and for the imputed sample under each of the three imputation models. The median imputation has an odds ratio very similar to that of the unimputed beneficiaries (1.36 vs. 1.37). In contrast, we obtain a notably different odds ratio from the unimputed sample when we employ the least positive response (1.51) and a somewhat different odds ratio when we impute the most positive response (1.31).

Table 11 explores the calibration of the imputation models by examining mortality for each risk-score point value. The most positive response imputation does not calibrate well; imputed cases show higher mortality than members of the unimputed sample at the same score. When the missing cases are imputed as the least positive response option and the risk score is ≥ 3, then mortality is lower than unimputed cases with equivalent risk scores. If we used the most positive response imputation approach, then we would be assigning missing cases a risk score that overestimates mortality consistently for all individual scores ≥ 3, whereas least positive response imputation would be miscalibrated in the opposite direction consistently. The recommended median imputation approach yields the best calibration. This pattern of findings suggests that item missingness does not provide additional information about mortality risk beyond the information contained in nonmissing 2009 VES-HOS items, at least among those completing 80% or more of the HOS. This in turn suggests that item missingness for the HOS is "missing at random (MAR)," conditional on the other 2009 VES-HOS items, so that it would also meet the assumptions for multiple imputation as an alternative approach (Rubin).

Results of Comparison of Gender Differences in Predictive Validity of the 2009 VES-HOS Algorithm (Tables 12-17)

The results in this section and the remainder of the report include the full sample with median imputed cases (N=238,687).

As was shown in Table 1, in 2009 women made up 58% of the sample, but only 52% of overall deaths by the two-year follow up. The mean age for women in the sample (75.9 years) was significantly greater ($p<0.0001$) than for men (75.1 years). Consequently, the risk profile of the sample as determined by 2009 VES-HOS Risk Scores may differ by gender (not shown in the table).

Table 12 compares the frequencies of individual 2009 VES-HOS Risk Scores by gender. On average, women score 0.5 to 2 points higher than men on risk scores (varying slightly with the range of risk scores), largely because of the greater average age of women relative to men. If the 2009 VES-HOS algorithm overestimates mortality for women and underestimates mortality for men (main effect) or the relationship between risk score and mortality differs by gender (differential slopes), then gender would ideally be incorporated into the risk score model. Table 13 assesses such possibilities. It shows that the increase in odds of death with each one point increase in risk score is similar for men and women, even though (as shown in Table 14) women have slightly lower mortality risk at every score than men. Ideally, one would capture this slight difference in mortality risk while also preserving the easy-to-use integer (whole number) scoring

scheme. Unfortunately, any adjustment as large as one point would "overshoot" this small difference, resulting in a scoring that would be inferior to the version that did not incorporate gender (since the difference in mortality was much smaller than what would result from integer changes). All versions of the VES use the easy-to-use integer scoring system. Given the utility of the simple integer-based-scoring (Saliba, et al., 2001) and the fact that the best integer-based scoring does not adjust for gender, we recommend the 2009 VES-HOS algorithm without adjustment for gender.

We also investigated whether a gender adjustment is indicated within specific age groups. Tables 15-17 present results from an alternate model in Table 14 that classifies beneficiaries into six categories involving each combination of gender and three-category age (65-74, 75-84, and 85 or older). While we note that within each age stratification women had somewhat lower mortality at a given risk score than men, we recommend no gender adjustment be made in the 2009 VES-HOS. We believe that the very small improvements in accuracy that a non-integer adjustment by gender might provide do not justify the loss of simplicity that a departure from integer-based VES scoring would entail.

Chapter 5: Describing the Characteristics of Each of Three VES-HOS Risk Groups

In the original VES-13 and the 2005 VES-HOS scoring approaches, two risk score categories were identified: low (0-2 points) and high (≥3 points), with the high-risk group facing significantly higher rate of death (in VES-13 and 2005 VES-HOS) and functional decline (VES-13 only) than the low-risk group. In the update and extension of the 2009 VES-HOS, we identify three risk score groups to provide more specific information: low (0-2 points), moderate (3-6 points), and high (≥7 points). This chapter describes the socio-demographic characteristics of seniors in each of the three risk groups, and the source of risk score points that tend to place seniors in each risk group.

Findings

We find that 38% of the 2009 HOS sample falls in the low-risk group (0-2 points), 29% in the moderate-risk group (3-6 points), and 33% in the high-risk group (7-10 points). Compared to low-risk seniors, moderate-risk seniors have twice the odds of dying in two years and high-risk seniors more than seven times the odds (Table 18).

Table 19 describes the socio-demographic characteristics of each risk group. Age is strongly related to risk group, since it is directly incorporated into the scoring algorithm. By definition, no one can be 85 years or older and be classified as low risk. Women and socio-economically disadvantaged seniors are more likely to be in the moderate and, especially, high risk categories than are men and those with higher SES. Black, widowed, divorced, separated, and less educated persons are more likely to be in the moderate and, especially, high-risk groups than other seniors. Tables 19 and 20 summarize the source and mean points from each source by risk category.

Comparison of Risk of Two-Year Mortality in Low, Moderate, and High-Risk Groups (Table 18)

The results demonstrate that the scoring system identifies members at increased risk of death over a two-year period, with higher scores generally identifying increasing risk of death. In both the original VES-13 and the 2005 VES-HOS 1.0 validation studies, individuals with scores ≥3 were considered vulnerable compared to those with lower scores <3 (Saliba, et al., 2001; Ng, et al., 2012). In this study, a score of ≥3 also identified a higher risk of two-year mortality than existed with a score of <3 (Table 18). Results demonstrate that the odds of dying within two years were 2.2 times higher for beneficiaries in the moderate-risk group relative to the low-risk group (those scoring <3). The high-risk group has 7.4 times the odds of dying within two years compared to the low-risk group.

Socio-Demographic Characteristics of Members of Vulnerable Groups (Table 19)

Table 19 compares the socio-demographic characteristics of beneficiaries in each of the three risk score groups. The top row provides the number of beneficiaries in each risk group. The proportion of the study sample with a score <3 was 38%, comparable to 39% in the 2005 HOS. The moderate-risk group contains 29% of the sample and the high-risk group includes 33%.

In Table 19, asterisks denote significant differences from the low-risk group ($p<0.05$, $p<0.01$, and $p<0.0001$), based on the results of logistic regression models predicting each row characteristic from indicators of being in the moderate and high-risk groups. Bolded values indicate at least a 10 percentage point difference from the low-risk group.

The low-risk group has the youngest age profile, with 64% of seniors <75 years of age and 13% aged 80 or older. In the moderate-risk group, 61% of seniors are <75 years of age and 21% are aged 80 or older. In the high-risk group, 26% are <75 years of age and 48% are aged 80 or older (with more than half of these greater than 85 years of age). Age is different from the other socio-demographic characteristics we examined because it is the only one that is directly incorporated into the risk score algorithm. By definition, no one can be 85 years or older and be classified as low risk. The ratio of women to men increases as risk group becomes more severe, which may be due entirely to women's greater longevity and thus greater mean age. Women represent 54% of low-risk beneficiaries, 57% of the moderate-risk group, and 62% of the high-risk group. The proportion of Black beneficiaries increases across low to high risk groups, while the proportion of White beneficiaries declines in higher risk groups.

Consistent with literature on marital status and health, (Rendall, Weden, Favreault, & Waldron, 2011; Schone & Weinick, 1998; Seeman, 2000) there are strong differences in risk group membership by marital status. Married beneficiaries represent almost two-thirds of the low-risk group compared to 45% of the high-risk group. By contrast, the widowed represent fewer than one in five low-risk beneficiaries compared to nearly two in five high-risk beneficiaries. The patterns by marital status are in part associated with the older age of widowed versus married beneficiaries.

Table 19 also shows results consistent with literature establishing a positive association of SES (as measured by education and annual household income) with health (Kitagawa & Hauser, 1973; Marmot & Wilkinson, 1999; House, Lepkowski, Kinney, Mero, Kessler, & Herzog, 1994). Among beneficiaries who did not graduate high school, very few are in the low-risk category and nearly half are in the high-risk category. This may be partially associated with age cohorts, since educational attainment of successive birth cohorts rose steadily through the 20th century. There is a strong association of risk category with household income, with one in four high risk beneficiaries on Medicaid, compared to only 7% of low risk beneficiaries on Medicaid.

Sources of VES-HOS Points by Risk Group (Tables 20-21)

Table 20 shows the sources of 2009 VES-HOS points by risk group. Age is the primary source of points in the low-risk group, with about one in three low-risk seniors receiving points for being 75-84 years of age. Age also distinguishes the high-risk group from the other two groups, with 27% of high-risk seniors receiving 3 points for being aged 85 and older (compared to 8% of moderate-risk seniors and, by definition, 0% of the low-risk seniors). Self-rated health plays a role in distinguishing the high-risk group from the other two groups, with 71% receiving any self-rated health points, compared to 17% of moderate and 5% of low-risk seniors; nonetheless, self-rated health is not a primary source of points for any group. Physical activity distinguishes each of the three groups. Of the low-risk group, 90% receive 0 physical activities points and 1% receives 2 such points. By contrast, 33% of the moderate-risk group receive 0 physical activities points and 19% receive 2 such points, and 3% of the high-risk group receive 0 physical activities

points and 84% receive 2 such points. Physical condition distinguishes the moderate and high-risk groups from the low-risk group, is the largest source of points for both moderate and high-risk groups, and accounts for two-thirds of all points in the moderate-risk group.

Chapter 6: Discussion

In order to improve performance, MAOs seek ways to identify at-risk patients who could benefit from outreach and intervention. The HOS contains important information that could help to determine which approaches are most likely to be effective. We use data from the 2009 HOS to update a predictive model to identify at-risk seniors. With this model, we are able to provide MAOs with an easily-implemented screening tool to identify vulnerable beneficiaries.

One of the most widely used and cited instruments for identifying vulnerable elders is the VES-13 (Saliba, et al., 2001). Ng et al. (2012) adapted the VES-13 by matching items from the 2005 HOS 1.0 to analogous items in the VES-13 to identify MAO beneficiaries aged 65 and older at risk of dying within two years. Because of changes in the HOS since 2005, the 2005 version of the VES-HOS could no longer be employed, and it was unclear whether a modified VES-HOS scoring could be reliably applied to the HOS 2.0 instrument to identify vulnerable seniors.

We investigated whether the HOS 2.0 survey can be used to accurately identify at-risk elders. Because the 2012 HOS instrument is the same as the 2009 instrument (with the exception of changes to how the height and weight items are collected), our results will allow MAOs to apply the updated VES-HOS algorithm to their HOS data collected with the HOS 2.0 to identify at-risk members.

The updated 2009 VES-HOS Risk Score has four components: age, self-rated health, physical activities and physical condition. The content and wording of age and self-rated health were virtually unchanged from the HOS 1.0 to 2.0 instrument, but the content of the physical activities and physical condition items changed substantially. Of particular concern were the changes to the physical activity items: the 2009 HOS 2.0 contained no corresponding questions with similar content for three of the four physical activity items used in the 2005 HOS 1.0. One (of five) physical condition items has been dropped from the HOS since 2005. In other cases, there were slight changes to question wording and/or response categories. Despite these differences, we were able to update and validate the proposed scoring using 2009 HOS 2.0. Despite the changes in items used in the scoring, the scoring is as predictive of two-year mortality with the 2009 HOS baseline data as it was using the 2005 HOS baseline data. We conclude that the 2009 VES-HOS algorithm, and by extension the original VES-13 approach, are robust ones that can be adapted for use with a range of health studies.

In addition to showing that the 2005 VES-HOS scoring could be modified to accommodate substantial changes to physical activity items and still distinguish vulnerable groups of seniors as well as the original scoring algorithm, we were able to extend the scoring approach in two important ways. First, we developed an easy to understand and implement imputation approach that allows us to retain respondents who were eligible to receive a risk score except for 2009 VES-HOS item non-response. When the median imputed cases are added to the sample, the association between risk score and two-year mortality rate remains similar to that reported using the 2005 HOS baseline data. Second, we distinguished among three rather than two risk score groups as had been done previously. The original VES-13 and the 2005 VES-HOS 1.0 distinguished between low (<3 points) and high (≥3 points) risk groups. We identified three distinct, meaningful risk groups that are approximately equal in size: low (0-2 points,

representing 38% of the sample), moderate (3-6 points, with 29% of the sample), and high (7-10 points, with 33% of the sample). The two-year mortality rates of these groups monotonically increase; relative to low-risk group, the odds ratio of dying within two years is 2.2 for the moderate-risk group and 7.4 for the high-risk group.

Findings and Implications

We modified the 2005 VES-HOS algorithm to account for changes within the 2009 HOS 2.0 survey items, and conducted three analyses to validate the updated 2009 VES-HOS scoring. We compared the distribution of the individual items used in the HOS 2.0 with those used in the HOS 1.0. For all items, including the substitute items, the distributions were similar. In the case of the substitute items, we selected cut offs that yielded as close a distribution to the HOS 1.0 item being replaced as possible. An additional validation involved a comparison of the 2005 and 2009 unimputed baseline data risk score distributions. These distributions were also similar. Finally, we compared the relationship between two-year mortality and the 2005 and 2009 VES-HOS Risk Scores and found that this relationship was very similar for both scoring algorithms.

Since the three validation tests indicated that the risk scores using 2009 HOS 2.0 behaved similarly to those reported by Ng et al. (2012) using 2005 VES-HOS, we conclude that the VES-HOS 2.0 update is validated and can be used for identifying vulnerable seniors by other users. That we were able to update the Ng et al. (2012) 2005 VES-HOS for use with the 2009 HOS 2.0 suggests that future changes to the HOS can be incorporated into the VES-HOS algorithm.

We investigated three possible extensions of the VES-HOS algorithm. The first related to the treatment of the 16% of the sample which is not assigned a risk score because of missing VES-HOS items. We investigated three imputation approaches: most positive response, least positive response, and median response imputation. The median imputation approach yielded an odds ratio associated with a one-unit increase in risk score (1.36) that most closely approximated that observed in the unimputed sample (1.37), and resulted in risk scores that were well-calibrated to observed mortality. We recommend that CMS and other users employ median imputation of missing values to retain seniors with any missing VES-HOS items. Doing so allows for a substantial increase in the sample for which a risk score can be assigned. This will be especially important for smaller MAOs that otherwise might have too few beneficiaries to allow for reliable measurement of the vulnerability of their senior members. When the tool is used to screen for risk of mortality, such imputation is also important for identifying as many vulnerable seniors as possible.

In the second extension we examined the value of explicitly incorporating gender into the VES-HOS scoring. Neither the original VES-13 nor the 2005 VES-HOS incorporated gender into the calculation of risk scores. Earlier research explored whether gender added to the predictive ability of the original VES-13 and concluded that it does not. While we note that women may have slightly lower mortality at a given risk score than men, we recommend no gender adjustment be made in the 2009 VES-HOS because we believe that the very small improvements in accuracy that a non-integer adjustment by gender might provide does not justify the loss of simplicity that a departure from integer-based VES scoring would entail.

The third extension we developed was the expansion of the number of risk score groups from two to three. We identified three risk groups: low (0-2 points), moderate (3-6 points), and high (7-10 points). We found that 38% of persons aged 65 and older are at low risk (0-2 points), 29% are moderate risk (3-6 points), and 33% are high risk (7-10 points) and that mortality differs markedly across these three groups. Our report described the socio-demographic characteristics of each of the three risk groups. Compared to the moderate and high-risk groups, the socio-demographic profile of the low-risk group is young (64% are <75 years of age), more often male, less often Black, more often married, and higher in SES (as measured by education, annual household income, and Medicaid status).

We also examined the sources of the 2009 VES-HOS points by risk group. Age is the primary source of points in the low-risk group and plays a role in distinguishing the high-risk group from the other two. Self-rated health plays a role in distinguishing the high-risk group from the other two groups, though it is not a primary source of points for any group. Physical activity distinguishes each of the three groups. Physical condition distinguishes the moderate and high-risk groups from the low-risk group, is the largest source of points for both moderate and high-risk groups, and accounts for two-thirds of all points in the moderate-risk group.

The overrepresentation of women, racial/ethnic minorities, and those with low SES in the highest-risk group emphasizes the need for continued attention to these beneficiaries. These improvements and innovations improve the ability of the 2009 VES-HOS algorithm to allow CMS and MAOs to identify the most vulnerable beneficiaries and to address their health care needs. MAOs can use the updated VES-HOS for use with their HOS data that was collected with the HOS 2.0 to assess the degree of mortality risk within its older patient population as well as to target specific members for quality improvement intervention. Because the scoring items used in the HOS 2.0 instrument remain the same through 2012, MAOs can also use the most recently available 2012 data to identify vulnerable seniors. The model is conceptually robust to changes in the specific survey items asked and can be used in the presence of missing information which extends its usefulness as, for example, a screening tool for health care practitioners in a variety of organizational settings.

Conclusions

An updated version of the VES-13 Risk Score, that involves alternative survey data and items, performed as well predicting older adult mortality as the original VES-13 and the modified 2005 VES-HOS that were based on different samples and survey items. We effectively extended the updated 2009 VES-HOS algorithm to account for missing data, a significant concern in surveys of older patients, and we identified three (rather than two) large and distinct risk groups.

Limitations

The HOS involves a two-year follow up. Prior work using the original VES-13 shows that the VES-13 is predictive of both mortality and functional decline within both shorter and longer intervals than two-years. However, we do not know the extent to which the 2009 VES-HOS algorithm predicts mortality within a shorter or longer time period, nor do we know the extent to which it predicts elders at risk of functional decline among survivors.

The HOS is restricted to the MAO population; therefore, we cannot be certain if the results would apply to a more general Medicare beneficiary population without further validation of the VES-HOS in a different population. However, we note that the original VES-13 was validated using a 1993 national survey of community-dwelling Medicare beneficiaries, prior to the advent of MAOs.

Future Work

Future work can also use multivariate analyses to distinguish the influence of age and the role of race/ethnicity from the direct roles of SES and marital status, in order to clarify how demographic patterns of VES-HOS Risk Scores are produced.

Since we were able to validate the 2005 VES-HOS Risk Score using the 2009-2011 HOS Cohort 12 Merged Baseline and Follow Up data, future work can use these risk scores to address substantive questions. Examples of such analyses include examining the distribution of moderate and high-risk beneficiaries across contracts and developing feedback to contracts about these distributions. In addition, we could examine how quality of care, as measured by the HOS, differs by risk status. This could be done both overall and by contract. Analyses of the latter type could determine the extent to which MAO performance for less vulnerable beneficiaries is predictive of their performance for more vulnerable beneficiaries. This would provide evidence of whether some MAOs appear to specialize in their care and have specific areas of organizational strengths or weaknesses that might inform quality improvement efforts.

References

Coleman, E., Wagner, E., Grothaus, L., Hecht, J., Savarino, J., & Buchner, D. (1998). Predicting hospitalization and functional decline in older health plan enrollees: Are administrative data as accurate as self-report? Journal of the American Geriatrics Society, 46 (4), 419-425.

De Leeuw, E., Joop, J., & Huisman, M. (2003). Prevention and treatment of item nonresponse. Journal of Official Statistics, 19 (2), 153-176.

Elliott, M., Edwards, C., Angeles, J., Hambarsoomians, K., & Hays, R. (2005). Patterns of unit and item nonresponse in the CAHPS® Hospital Survey. Health Services Research, 40 (6 Pt2), 2096-2119.

Haffer, S., & Bowen, S. (2004). Measuring and Improving Health Outcomes in Medicare: The Medicare HOS Program. Health Care Financing Review, 25 (40), 1-3.

House, J., Lepkowski, J., Kinney, A., Mero, R., Kessler, R., & Herzog, A. (1994). The Social Stratification of Aging and Health. Journal of Health and Social Behavior, 35 (3), 213-234.

Iqbal, S. R. (2007). The Veterans RAND 12 Item Health Survey 9 VR-12): What it is and How it is Used. Retrieved March 5, 2013, from www.chqoer.research.va.gov/docs/VR12.pdf

Kitagawa, E., & Hauser, P. (1973). Differential Mortality in the United States: A Study in Socioeconomic Epidemiology. Harvard University Press.

Marmot, M., & Wilkinson, R. (1999). Social Determinants of Health. Oxford University Press.

Min, L., Elliott, M., Wenger, N., & Saliba, D. (2006). Higher vulnerable elders survey scores predict death and functional decline in vulnerable older people. Journal of the American Geriatrics Society, 54 (3), 507-511.

NCQA. (2009). HEDIS Specifications for the Medicare Health Outcomes Survey (Vol. 6). Washington DC.

NCQA. (July 8, 2009). Quality Assurance Plan. HEDIS® 2009 Medicare Health Outcomes Survey. Washington DC.

Ng, J., Elliott, M., Scholle, S., Ahmed, K., Collins, R., & Bierman, A. (2012). Identifying older adults at high risk of mortality using the Medicare Health Outcomes Survey. The Journal of Ambulatory Care Management, 35 (4), 278-292.

Pacula, J., Boult, C., Reed, R., & Aliberti, E. (1997). Predictive validity of the Pra instrument among older recipients of managed care. Journal of the American Geriatrics Society, 45, 614-617.

Rendall, M., Weden, M., Favreault, M., & Waldron, H. (2011). The protective effect of marriage for survival: a review and update. Demography, 48 (2), 481-506.

Rubin, D. (n.d.). An Overview of Multiple Imputation. Retrieved from www.amstat.org/sections/srms/Proceedings/papers/1988 016.pdf

Saliba, D., Elliott, M., Rubenstein, L., Solomon, D., Young, R., Kambert, C., et al. (2001). The Vulnerable Elders Survey: A tool for identifying vulernable older people in the community. Journal of the American Geriatrics Society, 49, 1691-1699.

Schone, B., & Weinick, R. (1998). Health-related behaviors and the benefits of marriage for elderly persons. Gerontologist, 38 (5), 618-627.

Seeman, T. (2000). Health promoting effects of friends and family on health outcomes in older adults. AM J Health Promot, 14 (6), 362-370.

Vojta, C., Vojta, D., TenHave, T., Amaya, M., Lavizzo-Mourey, R., & Asch, D. (2001). Risk screening in a Medicare/Medicaid population. Journal of General Internal Medicine, 16 (8), 525-530.

Tables

Table 1: Baseline Demographic Category by Death Status at Two Year Follow Up, HOS *2009 Cohort 12 Baseline*

HOS Demographic	Alive at Follow Up[1] N (%)	Dead at Follow Up[1] N (%)
Total	220,818 (92.5%)	17,869 (7.5%)
Age[2]		
65-69	49,377 (22.4%)	1,758 (9.8%)
70-74	66,528 (30.1%)	2,989 (16.7%)
75-79	50,617 (22.9%)	3,557 (19.9%)
80-84	33,243 (15.1%)	3,894 (21.8%)
85+	21,053 (9.5%)	5,671 (31.7%)
Gender		
Male	92,780 (42.0%)	8,625 (48.3%)
Female	128,038 (58.0%)	9,244 (51.7%)
Race		
White	182,270 (82.5%)	14,890 (83.3%)
Black	23,045 (10.4%)	1,990 (11.1%)
Other/Unknown	15,503 (7.0%)	989 (5.5%)
Marital Status		
Married	123,412 (56.6%)	8,088 (46.0%)
Widowed	58,984 (27.1%)	6,824 (38.8%)
Divorced or Separated	27,757 (12.7%)	2,003 (11.4%)
Never Married	7,746 (3.6%)	686 (3.9%)
Education		
Did Not Graduate HS	59,731 (27.6%)	6,358 (36.5%)
High School Graduate	76,528 (35.3%)	5,977 (34.3%)
Some College	46,051 (21.2%)	3,057 (17.5%)
4 Year Degree or Beyond	34,454 (15.9%)	2,042 (11.7%)
Annual Household Income		
Less than $10,000	26,044 (13.1%)	2,754 (17.1%)
$10,000-$19,999	45,410 (22.8%)	4,278 (26.5%)
$20,000-$29,999	36,728 (18.4%)	3,012 (18.6%)
$30,000-$49,999	40,685 (20.4%)	2,542 (15.7%)
$50,000 or More	28,645 (14.4%)	1,406 (8.7%)
Don't Know	21,878 (11.0%)	2,165 (13.4%)
Medicaid Status		
Medicaid	33,294 (15.1%)	4,089 (22.9%)
Non-Medicaid	187,524 (84.9%)	13,780 (77.1%)

[1]Column percentages shown. Missing values for marital status, education, and annual household income.
[2]Risk points for increasing age are awarded in the VES-13 Based Risk Scoring

Table 2: Comparison of 13-Item Vulnerable Elder Survey (VES-13) Predictor Variables and Proxy Variables Used in the HOS *2005 Cohort 8 Baseline* and *2009 Cohort 12 Baseline*

VES-13 Predictor Variables	Proxy 2005 HOS Predictor Variables	Proxy 2009 Baseline HOS Predictor Variables
Age	Age	Age
Self-Rated Health	Self-Rated Health	Self-Rated Health
Limitations in Stooping, Kneeling, Bending	Limitations in Bending, Kneeling, Stooping	Alternative Variable to Explore: *ADL Disability in Dressing*
Limitations in Lifting, or Carrying Objects Up to 10 Pounds	Limitations in: 1. Lifting Groceries *or* 2. Vigorous Activities (Running, Lifting Heavy Objects, Participating in Strenuous Sports)	Alternative Variables to Explore: *ADL Disability in Eating* *Accomplished less than would have liked because of physical health*
Limitations in Reaching or Extending Arms Above Shoulder Level	No corresponding variable	No corresponding variable
Limitations in Writing, Handling, or Grasping Small Objects	No corresponding variable	No corresponding variable
Limitations Walking a Quarter Mile	Limitations in Walking Several Blocks	Alternate Variables to Explore: *Limitations in Climbing Several Flights of Stairs*
Limitations Performing Heavy Housework	Limitations in Moderate Activities (Pushing a Vacuum Cleaner, Moving a Table, Bowling, Playing Golf)	Limitations in Moderate Activities (Pushing a Vacuum Cleaner, Moving a Table, Bowling, Playing Golf)
IADL Disability in Shopping	Limitations in Lifting Groceries	Alternate Variable to Examine: *Physical or Mental Health Limits Doing Usual Activities (e.g., Self-Care, Work, or Recreation)* *Number of days in past 30 days physical health not good*
IADL Disability in Managing Money	1. Depressed Mood in Past Year 2. Emotional Health Limits Time Spent on Work/ Activities	1. Depressed Mood in Past Year 2. Emotional Problems Limit How Carefully Work/Activities Are Done
ADL Disability in Walking Across the Room	ADL Disability in Walking	ADL Disability in Walking
IADL Disability in Doing Light Housework	Physical Health Limits Time Spent on Work/Activities	Physical Health Limits Kind of Work/Activities
ADL Disability in Bathing or Showering	ADL Disability in Bathing	ADL Disability in Bathing

IADL is Instrumental Activity of Daily Living; ADL is Activity of Daily Living. No color=2005 variables are available in 2009.

Table 3: Complete Questions Used in the 13-Item Vulnerable Elder Survey (VES-13) and Survey Items Used in the HOS *2005 Cohort 8 Baseline* and *2009 Cohort 12 Baseline* with Risk Scoring Approach

VES-13 Items	Proxy 2005 HOS Items	Proxy 2009 HOS Items	VES-13 Based Risk Scoring (0-10 scale)
1. Age (in years)	Age (in years)	Age (in years)	1 point for age 75-84; 3 points for age ≥85.
2. In general, compared to other people your age, would you say that your health is [Excellent; Very good; Good; Fair*; Poor*]	In general, would you say your health is [Excellent; Very good; Good; Fair*; Poor*]	In general, would you say your health is [Excellent; Very good; Good; Fair*; Poor*]	1 point for fair *or* poor.
3. How much difficulty, on average, do you have with the following physical activities? [No difficulty; A little difficulty; Some difficulty; A lot of difficulty*; Unable to do*]	The following items are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much? [No, not limited at all; Yes, limited a little; Yes, limited a lot*]	The following items are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much? [Yes, limited a lot*; Yes, limited a little; No, not limited at all]	1 point for each * response in 3a–3f. Maximum of 2 points.
a. Stooping, crouching, kneeling	Bending, kneeling, or stooping	No corresponding question; but considered difficulty getting in/out of chairs. Because of health or physical problem, do you have any difficulty doing the following activities without special equipment or help from another person? [no, I do not have difficulty; Yes, I have difficulty*; I am unable to do this activity*]	
b. Lifting, or carrying objects up to 10 pounds	Lifting or carrying groceries *or* Vigorous activities, such as running, lifting heavy objects, participating in strenuous sports	No corresponding question; but consider accomplished less than would have liked due to physical problems [No, none of the time; Yes, a little of the time; Yes, some of the time*; Yes, most of the time*; Yes, all of the time*.	
c. Reaching or extending arms above shoulder level	No corresponding question		

Table 3: Complete Questions Used in the 13-Item Vulnerable Elder Survey (VES-13) and Survey Items Used in the HOS *2005 Cohort 8 Baseline* and *2009 Cohort 12 Baseline* with Risk Scoring Approach

d. Writing, handling, or grasping small objects	No corresponding question		
e. Walking a quarter mile	Walking several blocks	No corresponding questions; but consider: Climbing several flights of stairs. Does your health now limit you in these activities? If so, how much? [Yes, limited a lot*; Yes, limited a little; No, not limited at all]	
f. Heavy housework such as scrubbing floors or washing windows	Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf	Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf	
4. Because of your health or a physical condition, do you have any difficulty [No; Yes and does not get help; Yes and get help*; Don't do because of health*; Don't do].		Because of health or physical problem, do you have any difficulty doing the following activities without special equipment or help from another person? [no, I do not have difficulty; Yes, I have difficulty*; I am unable to do this activity*]	<u>4 points</u> for one or more * responses in 4a–4e.
a. Shopping for personal items	Does your health now limit you in: Lifting or carrying groceries? If so, how much? [No, not limited at all; Yes, limited a little; Yes, limited a lot*]	No corresponding question, but consider: For how many of past 30 days was physical health not good [zero days, 1-13 days, 14-30 days*]	

Table 3: Complete Questions Used in the 13-Item Vulnerable Elder Survey (VES-13) and Survey Items Used in the HOS *2005 Cohort 8 Baseline* and *2009 Cohort 12 Baseline* with Risk Scoring Approach

b. Managing money	Two alternate questions examined: 1. In the past year, have you felt depressed or sad much of the time? [No; Yes*] 2. During the past 4 weeks, have you: Cut down on the amount of time you spent on work or other activities as a result of any emotional problems? [No; Yes*]	1. In the past year, have you felt depressed or sad much of the time? [No; Yes*] 2. During the past 4 weeks, have you accomplished less than you would like with work or other regular daily activities due to emotional problems? [no, none of the time; yes, a little of the time; yes, some of the time*; yes, most of the time*; yes, all of the time*]
c. Walking across the room	Because of a health or physical problem, do have any difficulty with the following activity: Walking? [No, I do not have difficulty; Yes, I have difficulty*; I am unable to do this activity*]	Walking
d. Doing light housework	During the past 4 weeks, have you: Cut down on the amount of time you spent on work or other activities as a result of your physical health? [No; Yes*]	During the past 4 weeks, has physical health limited the kind of work or other activities? [No, none of the time; Yes, a little of the time; Yes, Some of the time*; Yes, most of the time*; Yes, all of the time*]
e. Bathing or showering	Because of a health or physical problem, do have any difficulty with the following activity: Bathing? [No, I do not have difficulty; Yes, I have difficulty*; I am unable to do this activity*]	Bathing

* indicates corresponding VES-13 and HOS responses most similar in wording and content, used for scoring purposes. The scoring is also described in Saliba et al. (2001) and on the VES-13 website (RAND).

Table 4: Distribution of Percentages and Means of VES-HOS Risk Score Items from 2005 HOS 1.0 and 2009 HOS 2.0 and Proposed Substitutes, HOS *2005 Cohort 8 Baseline* and *2009 Cohort 12 Baseline*

Item	2005 HOS Percentage or Mean	2009 HOS Percentage or Mean
Age in years (mean)	76.1 (SD=6.7)	75.3 (SD=6.6)
Age Groups		
65 to <75	48.6%	52.1%
75 to <85	40.4%	37.6%
85 and above	11.0%	10.3%
Self-Rated Health: Fair or Poor	30.5%	28.7%
Limitations in Physical Function		
Bending, kneeling or stooping	24.6%	--
Getting in/out of chairs		
No, I do not have difficulty	--	77.6%
Yes, I have difficulty	--	20.8%
I am unable to do this activity	--	1.6%
Lifting groceries *or* vigorous activities	52.7%	--
Accomplished less than would have liked due to physical problems		
No, none of the time	--	33.0%
Yes, a little of the time	--	22.2%
Yes, some of the time	--	24.1%
Yes, most of the time	--	13.4%
Yes, all of the time	--	7.3%
Limitations in walking several blocks: limited a lot	55.7%	--
Climbing several flights of stairs (2009)		
Yes, limited a lot	--	27.9%
Yes, limited a little	--	36.2%
No, not limited at all	--	35.9%
Moderate activities: limited a lot	22.6%	21.0%
Lifting groceries	16.2%	--
For how many days of past 30 days was physical health not good		
Zero days	--	56.5%
1-13 days	--	23.6%
14-30 days	--	19.9%
Presence of depressed mood much of the past year: Yes	12.5%	13.7%

Table 4: Distribution of Percentages and Means of VES-HOS Risk Score Items from 2005 HOS 1.0 and 2009 HOS 2.0 and Proposed Substitutes, HOS *2005 Cohort 8 Baseline* and *2009 Cohort 12 Baseline*

Cut down amount of time on work/activities because of emotional health: Yes (2005)	22.4%	--
Emotional Problems Limit How Carefully Work/Activities Are Done		
No, none of the time	--	60.6%
Yes, a little of the time	--	17.0%
Yes, some of the time	--	13.0%
Yes, most of the time	--	6.0%
Yes, all of the time	--	3.4%
Cut down amount of time on work/activities because of physical health: Yes	37.9%	--
During the past 4 weeks, has physical health limited the kind of work or other activities		
No, none of the time	--	34.8%
Yes, a little of the time	--	20.9%
Yes, some of the time	--	23.1%
Yes, most of the time	--	13.2%
Yes, all of the time	--	8.1
ADL Difficulty—Have difficulty or unable to do the following:		
Walking	37.9%	--
No, I do not have difficulty	--	68.4%
Yes, I have difficulty	--	28.6%
I am unable to do this activity	--	3.0%
Bathing	15.5%	--
No, I do not have difficulty	--	85.2%
Yes, I have difficulty	--	11.6%
I am unable to do this activity	--	3.2%

Table 5: Frequency of VES-HOS Risk Score Based on Presence of Survey Items from the HOS *2005 Cohort 8 Baseline* and *2009 Cohort 12 Baseline*

Score[a]	Percentage of 2005 HOS Sample with Score (n=97,258)	Percentage of 2009 HOS Sample with Score (n=199,992)
0	16.7	23.3
1	14.4	15.5
2	7.9 (38.9% scored <3)	2.8 (41.5% scored <3)
3 (cutoff)	4.3 (61.1% scored ≥3)	2.5 (58.5% scored ≥3)
4	3.7	3.8
5	6.0	8.7
6	11.8	12.4
7	17.3	15.7
8	10.9	9.2
9	3.2	2.7
10	3.9	3.5

[a] The VES-HOS risk scoring approach, yielding a 0-10 risk score with 10 representing highest risk of death, is described in Table 3. A score of ≥3 is considered vulnerable (Saliba, et al., 2001).

Table 6: Bivariate Logistic Regression Predicting Two-Year Mortality, Based on Single 0-10 Risk Score, HOS *2005 Cohort 8 Baseline* and *2009 Cohort 12 Baseline*

	2005		2009	
Variable	**Odds Ratio**	**95% CI**	**Odds Ratio**	**95% CI**
1-point increase	1.36***	1.34-1.37	1.37***	1.36-1.38

**P* < 0.05; ***P* < 0.01; ****P* < 0.0001
CI indicates confidence interval.

Table 7: Percentage with VES-HOS Risk Score who Died, HOS *2005 Cohort 8 Baseline* and *2009 Cohort 12 Baseline*

Score[a]	2005	2009
0	1.5	1.5
1	2.6	3.2
2	4.0 (2.4% scoring <3 died)	4.7 (2.3% scoring <3 died)
3 (cutoff)	6.8 (11.0% scoring ≥3 died)	9.4 (10.5% scoring ≥3 died)
4	4.8	3.6
5	4.9	3.6
6	5.4	5.4
7	9.6	9.8
8	16.9	16.9
9	18.2	20.2
10	31.3	33.3

[a] The VES-HOS risk scoring approach, yielding a 0-10 risk score with 10 representing highest risk of death, is described in Table 3. A score of ≥3 is considered vulnerable (Saliba, et al., 2001).

Table 8: Two-Year Mortality Rate for Unimputed Sample and Missing Data Sample, HOS *2009 Cohort 12 Baseline*

	2009 HOS Unimputed Sample[a] (n=199,992)	2009 HOS Missing Data Sample[b] (n=38,695)
Two-year mortality	7.1%	9.5%

[a] Does not include cases with missing responses to predictor variables from 2009 HOS sample.
[b] Includes only cases with missing responses to predictor variables from 2009 HOS sample.

Table 9: Frequency of 2009 VES-HOS Risk Score, Based on Presence of Predictor Variables from the HOS *2009 Cohort 12 Baseline*

Score[a]	Percentage of 2009 HOS Unimputed Sample[b] with Score (n=199,992)	Percentage of 2009 HOS Imputed (Least Positive Response) Sample[c] with Score (n=38,695)	Percentage of 2009 HOS Imputed (Median) Sample[c] with Score (n=38,695)	Percentage of 2009 HOS Imputed (Most Positive Response) Sample[c] with Score (n=38,695)
0	23.3	0	7.8	11.7
1	15.5	4.3	8.7	13.6
2	2.8 (41.5% scored <3)	4.4 (8.7% scored <3)	2.9 (19.4% scored <3)	4.4 (29.7% scored <3)
3 (cutoff)	2.5 (58.5% scored ≥3)	1.2 (91.3% scored ≥3)	2.3 (80.6% scored ≥3	4.1 (70.4% scored ≥3)
4	3.8	6.9	5.7	4.2
5	8.7	13.0	13.3	9.5
6	12.4	17.9	17.4	14.6
7	15.7	25.5	20.5	18.3
8	9.2	15.9	12.3	11.3
9	2.7	4.8	4.6	4.1
10	3.5	6.1	4.5	4.3

[a] The 2009 VES-HOS risk scoring approach, yielding a 0-10 risk score with 10 representing highest risk of death, is described in Table 3. A score of ≥3 is considered vulnerable (Saliba, et al., 2001).
[b] Does not include cases with missing responses to predictor variables from 2009 HOS sample.
[c] Includes only cases with missing responses to predictor variables from 2009 HOS sample.

Table 10: Bivariate Logistic Regression Predicting Two-Year Mortality, Based on Single 0-10 Risk Score for Unimputed and Imputed *HOS 2009 Cohort 12 Baseline*

	Unimputed Sample[a]		Imputed (Least Positive Response) Sample[b]		Imputed (Median) Sample[b]		Imputed (Most Positive Response) Sample[b]	
Variable	Odds Ratio	95% CI	Odds Ratio	95% CI	Odds Ratio	95% CI	Odds Ratio	95% CI
1-point increase	1.37***	1.36-1.38	1.51***	1.47-1.54	1.36***	1.34-1.39	1.31***	1.29-1.33

*$P < 0.05$; **$P < 0.01$; ***$P < 0.0001$

CI indicates confidence interval.

[a] Does not include cases with missing responses to predictor variables from 2009 HOS sample.

[b] Includes only cases with missing responses to predictor variables from 2009 HOS sample.

Table 11: Percentage with 2009 VES-HOS Risk Score who Died for Unimputed and Imputed HOS *2009 Cohort 12 Baseline*

Score[a]	Unimputed Sample[b] (n=199,992)	Imputed (Least Positive Response) Sample[c] (n=38,695)	Imputed (Median) Sample[c] (n=38,695)	Imputed (Most Positive Response) Sample[c] (n=38,695)
0	1.5	-	1.9	1.8
1	3.2	2.0	4.3	4.3
2	4.7 (2.3% scoring <3 died)	4.3 (3.1% scoring <3 died)	5.3 (3.5% scoring <3 died)	5.9 (3.6% scoring <3 died)
3 (cutoff)	9.4 (10.5% scoring ≥3 died)	5.5 (10.1% scoring ≥3 died)	9.6 (11.1% scoring ≥3 died)	9.5 (12.0% scoring ≥3 died)
4	3.6	2.9	3.9	6.3
5	3.6	4.1	4.7	6.0
6	5.4	5.2	6.6	6.9
7	9.8	9.3	10.6	11.1
8	16.9	14.9	16.4	16.5
9	20.2	21.7	22.8	24.5
10	33.3	28.8	31.2	31.5

[a] The 2009 VES-HOS risk scoring approach, yielding a 0-10 risk score with 10 representing highest risk of death, is described in Table 3. A score of ≥3 is considered vulnerable (Saliba, et al., 2001).
[b] Does not include cases with missing responses to predictor variables from 2009 HOS sample.
[c] Includes only cases with missing responses to predictor variables from 2009 HOS sample.

Table 12 Frequency of 2009 VES-HOS Risk Score by Gender, Based on Presence of Predictor Variables from the HOS *2009 Cohort 12 Baseline*

Score[a]	Percentage of 2009 HOS Full Sample with Score (n=238,687)[b]	Percentage of Males in 2009 HOS Full Sample with Score (n=101,405)[b]	Percentage of Females in 2009 HOS Full Sample with Score (n=137,282)[b]
0	20.8	23.1	19.0
1	14.4	15.6	13.5
2	2.8 (38.0% scored <3)	2.9 (41.6% scored <3)	2.7 (35.2% scored <3)
3 (at-risk cutoff)	2.5 (62.0% scored ≥3)	2.4 (58.5% scored ≥3)	2.6 (64.8% scored ≥3)
4	4.1	4.1	4.1
5	9.4	9.8	9.2
6	13.2	13.1	13.4
7	16.5	15.4	17.3
8	9.7	8.8	10.4
9	3.0	2.2	3.5
10	3.6	2.7	4.3

[a] The VES-HOS risk scoring approach, yielding a 0-10 risk score with 10 representing highest risk of death, is described in Table 3. A score of ≥3 is considered vulnerable (Saliba, et al., 2001).

[b] Includes missing values for VES-HOS predictor variables that were imputed to median of non-missing values from 2009 HOS sample. The variable for gender had no missing values. Percentages may add to more than 100% due to rounding.

Table 13 Logistic Regression Models Predicting Two-Year Mortality for Each Point of Risk Score by HOS *2009 Cohort 12 Baseline* Full Sample, Male Sample, and Female Sample

	2009 HOS Full Sample (n=238,687)[a] Overall Risk for Death		2009 HOS Males Full Sample (n=101,405)[a] Overall Risk for Death		2009 HOS Females Full Sample (n=137,282)[a] Overall Risk for Death	
Point Value	**Odds Ratio**	**95% CI**	**Odds Ratio**	**95% CI**	**Odds Ratio**	**95% CI**
0	Ref	Ref	Ref	Ref	Ref	Ref
1	2.22***	2.03-2.44	2.21***	1.95-2.49	2.29***	1.98-2.65
2	3.32***	2.90-3.80	2.97***	2.47-3.58	3.91***	3.22-4.74
3	6.83***	6.10-7.65	6.17**	5.26-7.24	8.08***	6.85-9.52
4	2.51***	2.21-2.85	2.45***	2.06-2.92	2.72***	2.25-3.29
5	2.63***	2.38-2.90	2.75***	2.41-3.13	2.59***	2.22-3.02
6	3.94***	3.61-4.30	3.95***	3.52-4.43	4.19***	3.67-4.78
7	7.27***	6.71-7.87	7.52***	6.77-8.36	7.68***	6.79-8.68
8	13.34***	12.31-14.45	14.20***	12.76-15.81	14.01***	12.38-15.85
9	17.33***	15.80-19.01	18.12***	15.85-20.71	19.65***	17.17-22.49
10	32.10***	29.47-34.95	32.02***	28.35-36.15	37.51***	33.05-42.58

*$P < 0.05$; **$P < 0.01$; ***$P < 0.0001$

CI indicates confidence interval.

[a] Includes missing values for VES-HOS predictor variables that were imputed to median of non-missing values from 2009 HOS sample. The variable for gender had no missing values.

Table 14: Percentage with 2009 VES-HOS Risk Score who Died by Gender, HOS *2009 Cohort 12 Baseline*

Score[a]	Percentage of 2009 HOS Full Sample with Score (n=238,687)[b]	Percentage of Males in 2009 HOS Full Sample with Score (n=101,405)[b]	Percentage of Females in 2009 HOS Full Sample with Score (n=137,282)[b]
0	1.5	1.9	1.1
1	3.3	4.1	2.6
2	4.8 (2.4% scoring <3 died)	5.4 (2.9% scoring <3 died)	4.3 (1.9% scoring <3 died)
3 (at-risk cutoff)	9.4 (10.6% scoring ≥3 died)	10.6 (12.5% scoring ≥3 died)	8.6 (9.3% scoring ≥3 died)
4	3.7	4.5	3.1
5	3.8	5.0	2.9
6	5.6	7.1	4.6
7	9.9	12.6	8.2
8	16.8	21.5	14.0
9	20.8	25.8	18.5
10	32.8	38.1	30.3

[a] The 2009 VES-HOS risk scoring approach, yielding a 0-10 risk score with 10 representing highest risk of death, is described in Table 3. A score of ≥3 is considered vulnerable (Saliba ,et al, 2001[1]).

[b] Includes missing values for VES-HOS predictor variables that were imputed to median of non-missing values from 2009 HOS sample. The variable for gender had no missing values.

Table 15: Percentage with 2009 VES-HOS Risk Score who Died by Gender for Ages 65-74, HOS *2009 Cohort 12 Baseline*

Score[a]	Percentage Aged 65-74 in 2009 HOS Full Sample with Score (n=120,652)[b]	Percentage of Males Aged 65-74 in 2009 HOS Full Sample with Score (n=53,713)[b]	Percentage of Females Aged 65-74 in 2009 HOS Full Sample with Score (n=66,939)[b]
0	1.5	1.9	1.1
1	2.9	3.9	2.2
2	3.0 (1.7% scoring <3 died)	3.3 (2.1% scoring <3 died)	2.8 (1.3% scoring <3 died)
3 (at-risk cutoff)	5.0 (6.0% scoring ≥3 died)	9.6 (7.4% scoring ≥3 died)	2.0 (4.9% scoring ≥3 died)
4	2.4	3.1	1.8
5	3.1	4.0	2.4
6	5.3	6.7	4.4
7	10.3	13.1	8.4
8	na	na	na
9	na	na	na
10	na	na	na

[a] The 2009 VES-HOS risk scoring approach, yielding a 0-10 risk score with 10 representing highest risk of death, is described in Table 3. A score of ≥3 is considered vulnerable (Saliba, et al., 2001).

[b] Includes missing values for VES-HOS predictor variables that were imputed to median of non-missing values from 2009 HOS sample. The variable for gender had no missing values.

Table 16: Percentage with 2009 VES-HOS Risk Score who Died by Gender for Ages 75-84, HOS *2009 Cohort 12 Baseline*

Score[a]	Percentage Aged 75-84 in 2009 HOS Full Sample with Score (n=91,311)[b]	Percentage of Males Aged 75-84 in 2009 HOS Full Sample with Score (n=38,417)[b]	Percentage of Females Aged 75-84 in 2009 HOS Full Sample with Score (n=52,894)[b]
0	na	na	na
1	3.3	4.1	2.7
2	5.4 (3.7% scoring <3 died)	6.2 (4.4% scoring <3 died)	4.9 (3.0% scoring <3 died)
3 (at-risk cutoff)	5.5 (10.7% scoring ≥3 died)	5.8 (13.5% scoring ≥3 died)	5.2 (8.8% scoring ≥3 died)
4	9.0	9.5	8.6
5	5.0	7.1	3.4
6	6.0	7.6	4.9
7	9.3	12.1	7.5
8	17.5	22.5	14.3
9	na	na	na
10	na	na	na

[a] The 2009 VES-HOS risk scoring approach, yielding a 0-10 risk score with 10 representing highest risk of death, is described in Table 3. A score of ≥3 is considered vulnerable (Saliba, et al., 2001).

[b] Includes missing values for VES-HOS predictor variables that were imputed to median of non-missing values from 2009 HOS sample. The variable for gender had no missing values.

Table 17: Percentage with 2009 VES-HOS Risk Score who Died by Gender for Ages 85+, HOS *2009 Cohort 12 Baseline*

Score[a]	Percentage Aged 85+ in 2009 HOS Full Sample with Score (n=26,724)[b]	Percentage of Males Aged 85+ in 2009 HOS Full Sample with Score (n=9,275)[b]	Percentage of Females Aged 85+ in 2009 HOS Full Sample with Score (n=17,449)[b]
0	na	na	na
1	na	na	na
2	na	na	na
3 (at-risk cutoff)	11.0 (21.2% scoring ≥3 died)	12.4 (24.1% scoring ≥3 died)	10.1 (19.7% scoring ≥3 died)
4	13.5	17.4	11.4
5	13.6	17.2	11.8
6	13.3	9.1	15.8
7	12.3	13.1	11.8
8	13.4	15.9	11.8
9	20.8	25.8	18.5
10	32.8	38.1	30.3

[a] The 2009 VES-HOS risk scoring approach, yielding a 0-10 risk score with 10 representing highest risk of death, is described in Table 3. A score of ≥3 is considered vulnerable (Saliba, et al., 2001).

[b] Includes missing values for VES-HOS predictor variables that were imputed to median of non-missing values from 2009 HOS sample. The variable for gender had no missing values.

Table 18: Logistic Regression Models Predicting Two-Year Mortality Based on Risk Group, Moderate Risk Versus Low Risk and High Risk Versus Low Risk, HOS *2009 Cohort 12 Baseline*

	2009 HOS Full Sample (n=238,687)[b]	
	Overall Risk for Death	
Variable	**Odds Ratio**	**95% CI**
Moderate Risk[c] vs. Low Risk[d]	2.18†	2.06-2.30
High Risk[e] vs. Low Risk[d]	7.42†	7.09-7.78

* $P < 0.05$; § $P < 0.01$; † $P < 0.0001$

CI indicates confidence interval.

[b] Includes missing values for 2009 VES-HOS predictor variables that were imputed to median of non-missing values from 2009 HOS sample. The variable for gender had no missing values.

[c] 2009 VES-HOS Risk Score of 3-6

[d] 2009 VES-HOS Risk Score of 0-2

[e] 2009 VES-HOS Risk Score of 7-10

Table 19: Frequency[1] of 2009 VES-HOS Risk group by Demographic Category from the HOS *2009 Cohort 12 Baseline*

HOS Demographic	Low Risk (Score 0-2) N (%)	Moderate Risk (Score 3-6) N (%)	High Risk (Score 7-10) N (%)
Total	90,597 (100.0%)	69,826 (100.0%)	78,264 (100.0%)
Age[2]			
65-69	24,232 (26.8%)	17,546 (25.1%)***	**9,357 (12.0%)***
70-74	33,698 (37.2%)	24,926 (35.7%)***	**10,893 (13.9%)***
75-79	21,327 (23.5%)	12,686 (18.2%)***	20,161 (25.8%)***
80-84	11,340 (12.5%)	8,992 (12.9%)*	16,805 (21.5%)***
85+	0 (0.0%)	5,676 (8.1%)	**21,048** (26.9%)***
Gender			
Male	42,176 (46.6%)	29,716 (42.6%)***	29,513 (37.7%)***
Female	48,421 (53.5%)	40,110 (57.4%)	48,751 (62.3%)
Race			
White	79,246 (87.5%)	55,828 (80.0%)***	62,086 (79.3%)***
Black	6,478 (7.2%)	8,786 (12.6%)***	9,771 (12.5%)***
Other/Unknown	4,873 (5.4%)	5,212 (7.5%)***	6,407 (8.2%)***
Marital Status			
Married	58,897 (65.7%)	37,829 (55.0%)***	34,774 (45.1%)***
Widowed	17,249 (19.3%)	18,258 (26.5%)***	**30,301 (39.3%)***
Divorced or Separated	10,521 (11.7%)	9,930 (14.4%)***	9,309 (12.1%)***
Never Married	2,917 (3.3%)	2,823 (4.1%)***	2,692 (3.5%)***
Education			
Did Not Graduate HS	16,091 (18.0%)	20,080 (29.3%)***	**29,918 (39.1%)***
High School Graduate	31,834 (35.7%)	24,643 (36.0%)	26,028 (34.1%)***
Some College	21,729 (24.3%)	14,377 (21.0%)***	13,002 (17.0%)***
4 Year Degree or Beyond	19,630 (22.0%)	9,378 (13.7%)***	**7,488 (9.8%)***
Annual Household Income			
Less than $10,000	5,900 (7.3%)	9,075 (14.3%)***	**13,823 (19.3%)***
$10,000-$19,999	13,748 (17.0%)	15,082 (23.8%)***	20,858 (29.2%)***
$20,000-$29,999	14,692 (18.2%)	12,261 (19.3%)	12,787 (17.9%)***
$30,000-$49,999	20,565 (25.5%)	12,474 (19.7%)***	10,188 (14.3%)***
$50,000 or More	17,806 (22.1%)	7,374 (11.6%)***	**4,871 (6.8%)***
Don't Know	7,969 (9.9%)	7,136 (11.3%)	8,938 (12.5%)***
Medicaid Status			
Medicaid	6,529 (7.2%)	**11,258 (16.1%)***	**19,596 (25.0%)***
Non-Medicaid	84,068 (92.8%)	58,568 (83.9%)	58,668 (75.0%)

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.0001$

[1]Column percent shown with bolded values indicating ±10% compared to low risk group

[2]Risk points for increasing age are awarded in the VES-13 Based Risk Scoring

Table 20: Sources of 2009 VES-HOS Points by Risk group, HOS *2009 Cohort 12 Baseline*

Source of Points	Low Risk (Score 0-2) N (%)	Moderate Risk (Score 3-6) N (%)	High Risk (Score 7-10) N (%)	Total N (%)
Total	90,597 (100.0%)	69,826 (100.0%)	78,264 (100.0%)	238,687 (100.0%)
Points for Age	$\bar{x}$=0.4	$\bar{x}$=0.6	$\bar{x}$=1.3	$\bar{x}$=0.7
0	57,930 (63.9%)	42,472 (60.8%)	20,250 (25.9%)	120,652 (50.6%)
1	32,667 (36.1%)	21,678 (31.1%)	36,966 (47.2%)	91,311 (38.3%)
3	0 (0.0%)	5,676 (8.1%)	21,048 (26.9%)	26,724 (11.2%)
Points for Self-Rated Health	$\bar{x}$=0.1	$\bar{x}$=0.2	$\bar{x}$=0.7	$\bar{x}$=0.3
0	86,091 (95.0%)	57,984 (83.0%)	22,454 (28.7%)	166,529 (69.8%)
1	4,506 (5.0%)	11,842 (17.0%)	55,810 (71.3%)	72,158 (30.2%)
Points for Physical Activities	$\bar{x}$=0.1	$\bar{x}$=0.9	$\bar{x}$=1.8	$\bar{x}$=0.9
0	81,083 (89.5%)	23,344 (33.4%)	2,036 (2.6%)	106,463 (44.6%)
1	8,495 (9.4%)	33,543 (48.0%)	10,256 (13.1%)	52,294 (21.9%)
2	1,019 (1.1%)	12,939 (18.5%)	65,972 (84.3%)	79,930 (33.5%)
Points for Physical Condition	$\bar{x}$=0.0	$\bar{x}$= 3.6	$\bar{x}$=4.0	$\bar{x}$=2.4
0	90,597 (100.0%)	7,556 (10.8%)	0 (0.0%)	98,153 (41.1%)
4	0 (0.0%)	62,270 (89.2%)	78,264 (100.0%)	140,534 (58.9%)

Appendix 1

Table A1: Two-Year Mortality Rate for Beneficiaries by Number of Data Items Missing Across Predictor Variables, HOS *2009 Cohort 12 Baseline*

Number of Items Missing	2009 HOS Dead at Follow Up (N=17,869)
0	7.1% (n=14,180)
1	9.3% (n=2,404)
2	10.1% (n=767)
3	9.7% (n=326)
4	10.9% (n=94)
5	15.0% (n=43)
6	9.5% (n=50)
7	10.0% (n=4)
8	0.0% (n=0)
9	100.0% (n=1)