

# Associations Between Health-Related Quality of Life and Mortality in Older Adults

Derek S. Brown · William W. Thompson ·  
Matthew M. Zack · Sarah E. Arnold · John P. Barile

Published online: 5 November 2013  
© Society for Prevention Research (outside the USA) 2013

**Abstract** This study measures the use and relative importance of different measures of health-related quality of life (HRQOL) as predictors of mortality in a large sample of older US adults. We used Cox proportional hazards models to analyze the association between general self-reported health and three “healthy days” (HDs) measures of HRQOL and mortality at short-term (90-day) and long-term (2.5 years) follow-up. The data were from Cohorts 6 through 8 of the Medicare Health Outcomes Survey, a national sample of older adults who completed baseline surveys in 2003–2005. At the long term, reduced HRQOL in general health and all categories of the HDs were separately and significantly associated with greater mortality ( $P < 0.001$ ). In multivariate analysis of long-term mortality, at least one HD category remained significant for each measure, but the associations between mental health and mortality were inconsistent. For short-term mortality, the physical health measures had larger hazard ratios, but fewer categories were significant. Hazard ratios decreased over time for all measures of HRQOL except mental health. In conclusion, HRQOL measures were shown to be significant predictors of short- and long-term mortality,

further supporting their value in health surveillance and as markers of risk for targeted prevention efforts. Although all four measures of HRQOL significantly predicted mortality, general self-rated health and age were more important predictors than the HDs.

**Keywords** Mortality · Survival analysis · Quality of life · Medicare

## Introduction

As the population ages and life expectancy increases in the USA, living a longer high-quality life in good health is becoming more important to the public health community, the medical community, and individuals (Moriarty et al. 2005; Netuveli and Blane 2008; Sondik et al. 2010). Questionnaires used on a variety of surveys are used to monitor the concept of health-related quality of life (HRQOL), a multidimensional construct of self-rated physical and mental health. “Generic” HRQOL instruments, which include a wide range of health domains, are collected on a number of large population surveys. For example, the Centers for Disease Control and Prevention’s (CDC) four-item measure of HRQOL (HRQOL-4; CDC 2000) has been collected since 1993 on the Behavioral Risk Factor Surveillance System (<http://www.cdc.gov/brfss/>), for example, as reported in Moriarty et al. (2005) and Moriarty et al. (2003). This measure, and other measures of HRQOL, are also included as part of ongoing population surveys, such as the Medical Expenditure Panel Survey (<http://meps.ahrq.gov/mepsweb/>) and the National Health Interview survey (<http://www.cdc.gov/nchs/nhis.htm>). Through such ongoing surveys, HRQOL measures provide a way to monitor both the health and well-being of the population over time and the potential association with risk factors. Several generic HRQOL measures have previously been compared in studies such as those of Hanmer

---

D. S. Brown (✉)  
Brown School, Washington University in St. Louis, One Brookings  
Drive, Campus Box 1196, St. Louis, MO 63130, USA  
e-mail: [dbrown@brownschool.wustl.edu](mailto:dbrown@brownschool.wustl.edu)

W. W. Thompson · M. M. Zack  
Division of Behavioral Surveillance, Public Health Surveillance  
Program Office, Centers for Disease Control and Prevention, Atlanta,  
GA, USA

S. E. Arnold  
Public Health Economics Program, RTI International, Research  
Triangle Park, NC, USA

J. P. Barile  
Department of Psychology, University of Hawai‘i at Mānoa,  
Honolulu, HI, USA

et al. (2006), Fryback et al. (2007), and Cherepanov et al. (2010). References on HRQOL include Bowling (1991), Spilker (1995), and Fayers and Machin (2007).

An emerging body of research has shown that monitoring HRQOL in population surveys may have additional uses. HRQOL measures, and in particular general self-rated health (GSRH), have been shown to be associated with future adverse health events, such as hospitalization (Cavrini et al. 2012; DeSalvo et al. 2005) and mortality (DeSalvo et al. 2005; Tsai et al. 2007; Sargent-Cox et al. 2010), although most studies to date have used selected samples. If this relationship is generalizable to other measures of HRQOL and to more representative samples of older adults, then HRQOL may be considered as an important tool in prevention for monitoring increased risk of adverse health events. One study (Dominick et al. 2002) found a significant relationship between HRQOL-4 and 1-month and 1-year mortalities, but longer-term associations, the relative importance of the four measures when used together (Dominick et al. 2002; Sargent-Cox et al. 2010; DeSalvo et al. 2006a, b), and the generalizability of results are unknown. Understanding these factors is important for surveillance and risk adjustment since health care resources are often heavily focused on end-of-life care.

This study investigates the following research questions: Do the items in HRQOL-4 predict short- or longer-term mortality? How much information do the three HDs in HRQOL-4 contribute relative to GSRH in predicting mortality? How do new results from a national sample compare to previous estimates?

## Research Design and Methods

### Data

The data are from Cohorts 6–8 of the Medicare Health Outcomes Survey (HOS), an annual national survey of Medicare beneficiaries voluntarily enrolled in Medicare Advantage (MA) private health plans (we were limited to three cohorts because of resource constraints, and the questionnaire switched versions beginning in Cohort 9, so we selected earlier years to ensure consistency in the measures). All managed care plans with MA contracts are required to participate in the HOS, which is used mainly for quality improvement programs, reports, and monitoring and improving health outcomes. When the data were collected, about 13 % of the Medicare population was enrolled in MA plans (Gold et al. 2010).

MA plans include persons aged 65+ and a small number of persons aged <65 with disabilities. Surveys are mailed to a randomly selected cohort of MA plan enrollees each spring, with a follow-up survey 2 years later. Non-respondents are surveyed by telephone. For large plans, a simple random

sample of 1,000 members are surveyed; for smaller plans with fewer than 1,000 members, all members are surveyed. HOS eligibility requires only continuous enrollment in the same plan for at least 6 months and no end-stage renal disease. Approximately 160,000 individuals are sampled for each cohort, and about 105,000 respond to the complete survey in each cohort.

This study obtained an HOS limited data set including dates of death from the Centers for Medicare & Medicaid Services and an exemption from the investigators' institutional review boards.

### Sample

The analysis sample includes all individuals aged 65+ from HOS Cohorts 6–8 whose baseline surveys were collected between 2003 and 2005. HOS members under 65 were excluded because they are eligible for Medicare due to disability and represent a specialized population. Individuals were also excluded from the analysis if they had missing values for any of the HRQOL-4 variables, marital status, or health indicators ( $n=96,583$ ). Proxy respondents were excluded ( $n=38,473$ ) because perception of HRQOL varies by reporter (Albrecht and Devlieger 1999; Andresen et al. 2001; Ellis et al. 2003). The total analysis sample was 191,001.

### Dependent Variable

The primary dependent variable is mortality status. Short-term mortality is defined occurring within 90 days of interview; long-term mortality is the status within the maximum follow-up length, slightly more than 2.5 years (analysis of 30 days and 1 year is available upon request). The time between the baseline survey and an individual's last observation—either date of death or the end of follow-up—was used in all Cox survival models described below. The exact date of death on the HOS is identified from the Medicare Enrollment Database and Social Security Administration Master Death File.

### Health-Related Quality of Life

The primary independent variables are the core CDC Healthy Days measures (HRQOL-4), which have been included on the HOS since 2003. These items have undergone cognitive testing and have demonstrated content, construct, criterion, and predictive validity, test–retest reliability, and internal consistency (Andresen et al. 2001; Moriarty et al. 2003; CDC 2000). The items include GSRH, physically unhealthy days (PUDs), mentally unhealthy days (MUDs), and days with activity limitations (ALDs). GSRH is assessed on a five-point Likert scale (“excellent,” “good,” “fair,” “poor,” or “bad”). PUDs are defined by the question “Now thinking about your physical health, which includes physical illness

and injuries, for how many days during the past 30 days was your physical health not good?"; MUDs by "Now thinking about your mental health, which includes stress, depression, and problems with emotions, for how many days during the past 30 days was your mental health not good?"; and ALDs by "During the past 30 days, for about how many days did poor physical or mental health keep you from doing your usual activities, such as self-care, work, or recreation?"

In the analysis, GSRH was used as a categorical variable with "excellent" as the reference group. PUDs, MUDs, and ALDs were grouped into 0-, 1- to 10-, 11- to 20-, and 21- to 30-day outcomes, with 0 day as the reference group, following past precedent (Dominick et al. 2002).

### Other Independent Variables

All analyses adjusted for age, sex, race/ethnicity, marital status, household income categories, educational status, and Spanish survey language. In addition, a count of 12 self-reported chronic diseases (hypertension or high blood pressure; stroke; emphysema, asthma or chronic obstructive pulmonary disease; Crohn's disease, ulcerative colitis, or inflammatory bowel disease; arthritis of hip, knee, hand or wrist; diabetes, high blood sugar, or sugar in the urine; congestive heart failure; any cancer other than skin cancer; angina pectoris or coronary heart disease; a myocardial infarction or heart attack; sciatica; other heart conditions, such as problems with heart valves or rhythm), six activities of daily living (ADLs; 0–2, 0=no difficulty, 1=has difficulty, 2=unable to do the activity), back pain (1–5, 5=greatest pain), smoking frequency (never, sometimes, or every day), and indicators for treatment of breast, colon, lung, and prostate cancer were included in all models. Survey mode and MA plan type indicators were also included.

### Statistical Analysis

Cox proportional hazards analysis of the risk of mortality as a function of time to death since baseline, HRQOL, and other covariates were run as both bivariate and multivariate models. Cox models (Cox 1972) are a semi-parametric survival model for estimating the rate of mortality based on the time to event. Censored data, in which mortality is not observed, enter the model naturally. The time variable in all models was the number of days from baseline until death, if observed, or the censoring date. The Cox model assumes proportional hazards over time, implying that the hazard ratio for a given covariate does not change over time. This assumption was tested by interacting HRQOL variables with time.

### Relative Importance

Two approaches were used to assess the relative importance of the three HDs compared to GSRH for predicting mortality.

Predictive validity was assessed using the C-index (18), which ranges from 0.5 to 1.0, and reflects the proportion of paired outcomes in which differing survival times are correctly predicted by the model. To compare the relative predictive power of different factors, the C-index was estimated for several models: (1) age; (2) age, GSRH; (3) age, HDs; and (4) age, GSRH, HDs. DeSalvo et al. (2005) and Quinten et al. (2009) also used similar approaches to assess the performance of HRQOL for predicting mortality outcomes. In addition, the full model (age, GSRH, HDs) was run with and without each individual term to compare the difference between the full model and each reduced form.

Relative importance was also assessed for the same models by the proportion of explained variation (PEV; Schemper 1993; Schemper and Stare 1996; Heinze and Schemper 2003). PEV measures the amount of variation in mortality (0–100 %) attributable to independent variables relative to the total variation in mortality. The partial PEV for each factor of interest reflects the PEV with and without the factor of interest.

## Results

### Sample Characteristics

Among the 191,001 respondents in this analysis, mean age was 75 years, 10 % were non-White, 58 % were female, and 58 % were married (Table 1). Within 90 days of the baseline survey date, 718 respondents (0.4 %) died. Over the maximum observation time of slightly more than 2.5 years, 13,414 (7 %) had died. Table 2 shows responses for each of the HRQOL-4 items. Compared to all persons in the HOS sampling frame, including non-respondents, the analysis sample is slightly younger (among those aged 65 and older) and has fewer minorities, more married persons, and higher education levels. All of these factors are associated with conditioning on non-proxy respondents and non-missing data. A detailed table of comparisons is available from the authors upon request.

### Bivariate Analysis

At the maximum follow-up time, age was the strongest predictor of mortality, with an estimated hazard ratio (HR) of 1.09 ( $P < 0.0001$ ) at age 65 that increased with age (Table 3). The HRs for each category of the four HRQOL variables were also highly significant ( $P < 0.001$ ), were greater than 1, and increased with each decrement in HRQOL. For any level of HDs (1–10, 11–20, and 21–30), the HRs were smallest for MUDs and largest for ALDs, ranging from 1.23 for 1–10 MUDs to 3.65 for 21–30 ALDs. Most other covariates had the expected relationship with mortality, such as increased HRs for males, persons with more chronic diseases, for those with

**Table 1** Characteristics of the analysis sample

Characteristic	<i>n</i>	Percentage <sup>a</sup>
Age		
65–69 years	47,099	24.7
70–74	58,651	30.7
75–79	45,015	23.6
80 or older	40,236	21.1
Female	111,373	58.3
Male	79,628	41.7
Race/ethnicity		
White	171,053	89.6
Black/African American	13,387	7.0
Other race/ethnicity	6,561	3.4
Married	109,768	57.5
Divorced, separated, widowed, never married	81,233	42.5
Annual household income		
<\$5,000	4,304	2.3
\$5,000–\$9,999	13,316	7.0
\$10,000–\$19,999	46,511	24.4
\$20,000–\$29,999	36,405	19.1
\$30,000–\$39,999	22,068	11.6
\$40,000–\$49,999	13,068	6.8
\$50,000 or more	19,760	10.4
Income missing/do not know	35,569	18.6
Mean chronic disease score (0–12) <sup>a,b</sup>	191,001	2.44
Activities of daily living (ADL) score (0–12) <sup>a</sup>	191,001	0.87
Back pain score (1–5, 5=highest) <sup>a</sup>	191,001	1.88
Current smoking frequency		
Non-smoker	169,966	89.0
Sometimes	5,866	3.1
Every day	14,003	7.3
Do not know/missing	1,166	0.6
Cancer treatment indicator		
Colon cancer	1,857	1.0
Lung cancer	983	0.5
Breast cancer	3,558	1.9
Prostate cancer	5,504	2.9
“‘No” to 1 or more types and skipped 1 or more types	2,395	1.3
Education		
Less than high school	42,583	22.3
High school diploma	74,586	39.1
Some college	42,730	22.4
4-year college graduate or higher	30,119	15.8
Education missing	983	0.5
Metro (USDA urban–rural continuum ≤3)	174,062	91.1
Non-metro (USDA urban–rural continuum >3)	16,939	8.9
English language survey	190,424	99.7
Spanish language survey	577	0.3
Mail survey	159,310	83.4
Telephone survey	31,691	16.6

**Table 1** (continued)

Characteristic	<i>n</i>	Percentage <sup>a</sup>
Medicare only eligible	182,410	95.5
Dual eligible (Medicaid and Medicare)	8,591	4.5
Medicare Advantage Plan type		
Competitive medical plan	88,836	46.5
Health maintenance organization (HMO)	92,474	48.4
Other	9,691	5.1
Medicare Advantage Plan model		
Individual practice association	99,979	52.3
Group practice	70,405	36.9
Staff model	15,435	8.1
Other	5,182	2.7
Medicare Advantage Plan tax status		
Not for profit	81,589	42.7
For profit	106,535	55.8
Not available	2,877	1.5
Cohort 6	64,235	33.6
Cohort 7	61,137	32.0
Cohort 8	65,629	34.4

<sup>a</sup> Mean values instead of percentages shown, where indicated

<sup>b</sup> The chronic diseases are: hypertension or high blood pressure; stroke; emphysema, asthma, or chronic obstructive pulmonary disease; Crohn's disease, ulcerative colitis, or inflammatory bowel disease; arthritis of hip, knee, hand, or wrist; diabetes, high blood sugar, or sugar in the urine; congestive heart failure; any cancer other than skin cancer; angina pectoris or coronary heart disease; a myocardial infarction or heart attack; sciatica; other heart conditions, such as problems with heart valves or rhythm

ADLs, and smokers (all  $P < 0.0001$ ). Married respondents and persons with higher income (both  $P < 0.0001$ ) had decreased HRs. HRs were  $< 1.0$  for those responding in Spanish ( $P = 0.004$ ) and for races other than White and Black ( $P = 0.001$ ), and HRs exceeded 1.0 for telephone respondents ( $P < 0.0001$ ).

### Multivariate Results

For short-term (90-day) survival, HRs increased with worse categories of HRQOL for GSRH, PUDs, and ALDs. HRs for PUDs and ALDs were significant at 11–20 days ( $P < 0.05$ ) and 21–30 days ( $P < 0.001$ ), roughly 1.4 and 1.7 for each measure, with confidence intervals overlapping across comparable levels of the two variables (second and third columns of Table 3). HRs for “fair” and “poor” GSRH were statistically significant ( $P < 0.001$ ) and larger than HRs for any levels of PUDs or ALDs. Contrasting with the other measures and the bivariate results, the HR for the only significant level of MUDs was  $< 1$ , indicating a decreased risk of mortality ( $P < 0.05$ ) with 11–20 days of “not good” mental health in the past 30 days relative to no days of “not good” mental health. The relationship for other control variables was largely as expected, although fewer

**Table 2** HRQOL measured by Healthy Days

HRQOL variables	<i>n</i>	Percentage
General health		
Excellent	11,800	6.2
Very good	51,914	27.2
Good	82,758	43.3
Fair	37,670	19.7
Poor	6,859	3.6
Recent days of “not good” physical health		
0 day	120,491	63.1
1–10 days	39,624	20.8
11–20 days	12,450	6.5
21–30 days	18,436	9.7
Recent days of “not good” mental health		
0 day	146,718	76.8
1–10 days	30,150	15.8
11–20 days	7,844	4.1
21–30 days	6,289	3.3
Recent days of activity limitation		
0 day	150,905	79.0
1–10 days	22,204	11.6
11–20 days	8,466	4.4
21–30 days	9,426	4.9

of these variables were significantly associated with mortality relative to the bivariate analysis. However, back pain score is associated with decreased risk of mortality ( $P < 0.001$ ) compared to increased risk in the bivariate analysis.

Over the maximum follow-up, several additional relationships that were not significant over 90 days were statistically significant. HRQOL results were consistent with the 90-day period, although the HRs are closer to 1.0 in most cases except for MUDs, indicating that the proportional hazards assumption may not have been met. Likewise, worse HRQOL corresponded to larger HRs, except for MUDs. All included categories of GSRH were significantly associated with mortality ( $P < 0.05$  or better). All categories of ALDs were significant at  $P < 0.01$  or better, but only 21–30 days of PUDs was significant ( $P < 0.001$ ). At 90 days, greater MUDs were associated with a decreased risk of mortality, but only 21–30 days was statistically significant ( $P < 0.01$ ).

### Specification Tests

To test the proportional hazards assumption of the Cox model, time interactions with the HRQOL variables were included in an additional multivariate model of maximum follow-up length (Table 4). Except for “fair” GSRH, all interactions were statistically significant for the same HRQOL terms that were significant in the corresponding model without time interactions

in Table 3. Between short- and long-term follow-up, the combined HR for “poor” GSRH decreased from 7.1 to 3.5 ( $P < 0.05$ ). For PUDs, only 21–30 days was significantly associated with increased mortality ( $P < 0.05$ ), and the HR declined from 1.5 to 1.1 over the same period. The HRs for ALDs decreased from 1.6 to 1.1 for 11–20 days ( $P < 0.05$ ) and from 1.9 to 1.2 for 21–30 days ( $P < 0.01$ ). The only significant interaction for MUDs, 11–20 days ( $P < 0.05$ ), moved in the opposite direction over time, with HR increasing from 0.8 to 1.1 by 2.5 years. All other covariates in Table 3 were also included in these models. For the time interactions at 90 days only, interactions between GSRH and time were significant.

### Model Fit and Predictive Validity

Using the total C-index, the HRQOL variables added more information to total fit in the short-term model, while age added more information over the long-term model (Table 5, first panel). The added contribution of the four HRQOL variables ranged from 0.12 to 0.17 (19–27 % relative increase) at 90 days compared to only 0.07–0.08 (10–11 %) at 2.5 years. For overall cohorts and both follow-up lengths, GSRH alone added more to the fit than all three of the HD measures.

The relative contributions of age, GSRH, and the three HDs were assessed in terms of the average change in PEV or individual contributions to the full model C-index (Table 5, bottom panel). GSRH was the single largest predictor for the short-term model of 90 days and the second most important predictor in all cohorts at 2.5 years. Age was the most important predictor over the long-term model of about 2.5 years and the second most important predictor in the 90-day model. The relative importance of all three HD measures, compared to GSRH and age, was lower in both the short- and long-term models, whether assessed by PEV or C-index. In the three healthy days measures, the long-term results display a clear pattern in both the PEV and C-index results: ALDs are most predictive of long-term mortality, PUDs the next most predictive, and MUDs provide almost no information. In the short-term models, the pattern is not as consistent. The PEV shows ALDs to be most predictive of the three, while MUDs and PUDs vary in rank depending on the cohort. The C-index results show PUDs more predictive in Cohorts 7–8 and ALDs in Cohort 6, with MUDs always at the bottom.

### Discussion

This study provides new data describing the relationship between HRQOL and mortality by comparing the relationship with short- and long-term mortality. It is also the first to examine the association between healthy days and mortality in a large nationwide population, to assess all four HRQOL-4

**Table 3** Cox Proportional Hazards Analysis, 3 Models of Mortality

	Bivariate: maximum follow-up ( <i>N</i> =191,001, deaths=13,414)		Multivariate: mortality in 90 days ( <i>N</i> =191,001, deaths=717)		Multivariate: maximum follow-up ( <i>N</i> =191,001, deaths=13,414)	
	HR	95% CI	HR	95% CI	HR	95% CI
General health (reference: excellent)						
Very good	1.23***	1.10–1.38	1.00	0.56–1.78	1.13*	1.01–1.27
Good	2.10***	1.88–2.35	1.65	0.95–2.85	1.60***	1.43–1.79
Fair	4.45***	3.99–4.98	3.03***	1.73–5.30	2.52***	2.25–2.83
Poor	11.04***	9.83–12.41	7.36***	4.08–13.25	4.24***	3.73–4.82
Recent days of “not good” physical health (reference: 0 day)						
1–10 days	1.30***	1.24–1.36	1.04	0.82–1.32	1	0.95–1.05
11–20 days	2.13***	2.01–2.26	1.37*	1.01–1.86	1.07	1.00–1.16
21–30 days	3.20***	3.06–3.34	1.72***	1.32–2.23	1.22***	1.14–1.30
Recent days of “not good” mental health (reference: 0 day)						
1–10 days	1.23***	1.17–1.29	1.03	0.84–1.26	0.99	0.94–1.04
11–20 days	1.97***	1.85–2.11	0.68*	0.50–0.94	0.97	0.90–1.05
21–30 days	2.36***	2.20–2.53	0.75	0.56–1.01	0.89***	0.82–0.96
Recent days of activity limitation (reference: 0 day)						
1–10 days	1.49***	1.42–1.57	1.21	0.94–1.57	1.10***	1.04–1.17
11–20 days	2.34***	2.19–2.49	1.40*	1.03–1.90	1.20***	1.11–1.29
21–30 days	3.65***	3.47–3.84	1.69***	1.29–2.21	1.36***	1.26–1.46
Age	1.09***	1.09–1.10	1.06***	1.05–1.07	1.08***	1.08–1.09
Male	1.46***	1.41–1.51	1.67***	1.42–1.98	1.67***	1.60–1.73
Race/ethnicity (reference: White)						
Black/African-American	0.99	0.93–1.06	0.88	0.65–1.18	0.81***	0.75–0.87
Other race/ethnicity	0.85***	0.77–0.94	1.04	0.70–1.54	0.83***	0.75–0.92
Married	0.69***	0.67–0.72	0.91	0.77–1.07	0.84***	0.81–0.87
Household income	0.87***	0.86–0.88	0.99	0.93–1.05	0.98*	0.97–1.00
Income missing or do not know	0.59***	0.56–0.63	1.04	0.77–1.39	0.98	0.92–1.05
Chronic disease score (0–12, 12=highest)	1.24***	1.23–1.25	1.04	1.00–1.08	1.07***	1.06–1.08
Activities of daily living (ADL) score (0–12)	1.22***	1.22–1.23	1.10***	1.06–1.14	1.09***	1.08–1.10
Back pain score (1–5)	1.18***	1.16–1.19	0.82***	0.77–0.88	0.89***	0.88–0.90
Nonsmoker (reference)						
Sometimes smoke	1.56***	1.44–1.70	1	0.65–1.53	1.67***	1.53–1.81
Smoke daily	1.60***	1.52–1.69	1.62***	1.27–2.07	1.87***	1.77–1.98
Do not know/missing	1.13	0.91–1.39	1.21	0.54–2.72	1.03	0.83–1.27
Cancer treatment status						
Colon cancer	1.97***	1.76–2.20	1.47	0.94–2.29	1.48***	1.32–1.65
Lung cancer	7.74***	7.02–8.54	5.68***	4.20–7.67	4.87***	4.42–5.36
Breast cancer	1.22***	1.10–1.36	1.26	0.82–1.96	1.30***	1.17–1.45
Prostate cancer	1.75***	1.62–1.89	1.17	0.85–1.61	1.11**	1.03–1.20
“No” 1+ types, skipped 1+ types	1.81***	1.60–2.04	1.15	0.66–2.03	1.25***	1.10–1.41
Education (reference: less than high school)						
High school diploma	0.75***	0.72–0.78	1.09	0.91–1.31	1.03	0.99–1.08
Some college	0.72***	0.69–0.76	0.92	0.74–1.16	1.02	0.97–1.07
4-year college graduate or higher	0.62***	0.59–0.66	0.8	0.60–1.06	0.94	0.89–1.00
Education missing	0.85	0.68–1.07	0.86	0.32–2.33	0.91	0.73–1.15
Metro (USDA urban–rural continuum ≤3)	1.13***	1.07–1.21	1.08	0.82–1.41	1.13***	1.06–1.20

**Table 3** (continued)

	Bivariate: maximum follow-up ( <i>N</i> =191,001, deaths=13,414)		Multivariate: mortality in 90 days ( <i>N</i> =191,001, deaths=717)		Multivariate: maximum follow-up ( <i>N</i> =191,001, deaths=13,414)	
	HR	95% CI	HR	95% CI	HR	95% CI
Language of survey administration						
Spanish	0.54**	0.36–0.82	1.01	0.25–4.13	0.60*	0.39–0.91
Mode of completion (baseline survey)						
Telephone	1.57***	1.50–1.63	1.15	0.96–1.39	1.34***	1.28–1.40
Dual eligible (Medicaid & Medicare)	1.57***	1.47–1.68	1.05	0.78–1.42	1.09*	1.01–1.17
Competitive medical plan (reference)						
HMO	1.08***	1.04–1.12	1.02	0.88–1.19	1.06**	1.02–1.09
Other	1.14**	1.06–1.23	0.94	0.55–1.61	1.12	0.99–1.26
Individual practice association (reference)						
Group practice	0.99	0.95–1.02	1.03	0.88–1.21	0.98	0.94–1.01
Staff model	1.01	0.95–1.08	0.89	0.66–1.21	1.01	0.94–1.07
Other	1.02	0.92–1.13	1.33	0.71–2.52	0.98	0.85–1.14
Medicare Advantage plan (reference: not for profit)						
For profit	1.01	0.97–1.04	1.05	0.90–1.23	1.01	0.97–1.05
Not available	1.21**	1.06–1.37	1.3	0.69–2.44	1.16	1.00–1.36
Cohort 7 indicator	0.98	0.94–1.02	0.88	0.74–1.06	0.96	0.92–1.00
Cohort 8 indicator	1.03	0.99–1.07	0.84*	0.70–1.00	0.96*	0.92–1.00

*N* is the sample size  
 \**P*<0.05; \*\**P*<0.01; \*\*\**P*<0.001

measures in a single model, and to follow them for more than 1 year.

The bivariate results are consistent with the findings of Dominick et al. (2002), although HRs are smaller in this study for comparable models. The HRs <1.0 for Spanish surveys and other (non-White, non-Black) race were unexpected, but may reflect a crossover association with survival (Corti et al. 1999; Johnson 2000). The MA population is healthier, on average, than the general population (Atherly et al. 2005; Mello et al. 2003; Lied et al. 2003), which may exaggerate a crossover association if selection into MA plans varies by race/ethnicity.

Comparing the short (90-day) and long (about 2.5 years) multivariate models, the relationship between HRQOL and mortality weakens as time since measurement increases. Although more covariates and HRQOL categories were statistically significant over the longer follow-up period, this results from increased power and more deaths observed, not increased predictive power for the HDs, as shown in Table 5. An exception to this is that the GSRH measure is more predictive over the long-term model than the short-term model. The weaker long-term relationship with mortality for the three HD measures is not unexpected since these capture

HRQOL during the past 30 days while the follow-up period is substantially longer. Physical health (GSRH and PUDs) and ALDs had the expected relationship with mortality, but the associations between mental health and mortality were less consistent and, in some cases, were in the opposite direction than expected. The finding of no significant relationship between mental health and mortality after adjusting for other measures of HRQOL was consistent with several other studies (Sargent-Cox et al. 2010; Steptoe and Wardle 2011; DeSalvo et al. 2006a, b; Franks et al. 2003; Korten et al. 1999; Vogt et al. 1994). One study examining the association between mental health and mortality (Pratt 2009) found significant associations between poor mental health and mortality, but did not control for HRQOL measures, including GSRH.

Assessing predictive validity and relative importance, the finding of a much smaller (10–11 % vs. 19–27 %) impact of the HRQOL-4 variables at the long term than at the short term is expected because the HRQOL-4 questions are phrased in terms of the past 30 days. If only one HRQOL measure were available for mortality modeling—over any length of time—GSRH is by far the most predictive. Among the three HDs, ALDs are the largest predictor in the PEV analysis of long-term follow-up and also the most stable, based on the smaller

**Table 4** Combined hazard ratio from Cox proportional hazards of time-interacted model (maximum follow-up length)

HRQOL variables	Main coefficient	Interacted coefficient	Combined hazard ratio at time in days						
			30	90	180	365	548	720	913
General health (reference: excellent)									
Very good	0.053	0.012	1.10	1.11	1.12	1.13	1.13	1.14	1.14
Good	0.496	−0.004	1.62	1.61	1.61	1.60	1.60	1.60	1.60
Fair	1.301***	−0.062	2.97	2.78	2.66	2.55	2.48	2.44	2.41
Poor	2.652***	−0.205*	7.06	5.64	4.89	4.23	3.90	3.68	3.51
Recent days of “not good” physical health (reference: 0 day)									
1–10 days	−0.015	0.003	0.99	1.00	1.00	1.00	1.00	1.00	1.00
11–20 days	0.412	−0.056	1.25	1.17	1.13	1.08	1.06	1.04	1.03
21–30 days	0.741**	−0.091*	1.54	1.39	1.31	1.23	1.18	1.15	1.13
Recent days of “not good” mental health (reference: 0 day)									
1–10 days	0.123	−0.022	1.05	1.02	1.01	0.99	0.98	0.98	0.97
11–20 days	−0.627*	0.101*	0.75	0.84	0.90	0.97	1.01	1.04	1.06
21–30 days	−0.512	0.067	0.75	0.81	0.85	0.89	0.91	0.93	0.95
Recent days of activity limitation (reference: 0 day)									
−10 days	0.208	−0.018	1.16	1.13	1.12	1.10	1.10	1.09	1.09
11–20 days	0.809**	−0.105*	1.57	1.40	1.30	1.21	1.15	1.12	1.09
21–30 days	1.122***	−0.139***	1.92	1.64	1.49	1.35	1.28	1.23	1.19

All other covariates used in Table 3 were included in these models, but are not shown

\* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$

**Table 5** Total model fit and contribution of individual variables

Model	C-index: 90-day follow-up			C-index: maximum 2.5-year follow-up		
	Cohort 6	Cohort 7	Cohort 8	Cohort 6	Cohort 7	Cohort 8
Age	0.6067	0.6403	0.6369	0.6548	0.6582	0.6496
Age+GSRH	0.7626	0.7456	0.7974	0.7202	0.7296	0.7183
Age+HDs	0.7301	0.7407	0.7489	0.6988	0.7042	0.6921
Age, GSRH, HDs	0.7723	0.7647	0.8077	0.7249	0.7333	0.7220
Variable	Cohort 6		Cohort 7		Cohort 8	
	Partial PEV of factor (%)	Contribution to full model C-index	Partial PEV of factor (%)	Contribution to full model C-index	Partial PEV of factor (%)	Contribution to full model C-index
90-day follow-up						
Age	0.10	0.0206	0.12	0.0473	0.18	0.0281
GSRH	0.43	0.0422	0.19	0.0240	0.50	0.0588
PUDs	0.01	0.0015	0.04	0.0088	0.02	0.0045
MUDs	0.06	0.0013	0.01	0.0017	0.02	0.000
ALDs	0.14	0.0040	0.12	0.0020	0.03	0.0016
Maximum 2.5-year follow-up						
Age	2.08	0.0573	2.28	0.0541	1.82	0.0505
GSRH	1.20	0.0261	1.56	0.0291	1.33	0.0299
PUDs	0.06	0.0011	0.03	0.0008	0.02	0.0004
MUDs	0.00	0.0002	0.00	0.0000	0.02	0.0002
ALDs	0.10	0.0011	0.14	0.0009	0.11	0.0013

Contribution to full model C-index found as the difference between the full model C-index (age+GSRH+HDs) and a model without the factor of interest  
*GSRH* general self-rated health, *HDs* healthy days, *PUDs* physically unhealthy days, *MUDs* mentally unhealthy days, *ALDs* days with activity limitations, *PEV* proportion of explained variation



time interactions with this variable. Thus, ALDs appear to be the most important of the three HD measures for long-term prognostic assessment. This conclusion should be weighed against the fact that the overall contribution of ALDs was still small in the models. Adding further robustness to these results, the differences in the contributions of the different HRQOL measures to the change in the C-index and PEV are consistent with a previous study that used another measure of HRQOL (SF-36) and reported C-index statistics for the alternative models (DeSalvo et al. 2005). In addition, the large association with the GSRH question after adjusting for other demographics, comorbid conditions, and depression has been shown in one previous study (Sargent-Cox et al. 2010). Another study that examined the variance in HRQOL explained by GSRH found that it explained 48–59 % of the variance in one HRQOL scale, the -12V instrument (DeSalvo et al. 2006a, b).

These findings have several implications for the monitoring of population health. Because the greatest level of predictive information is added by the GSRH item in HRQOL-4, the most parsimonious way to monitor for increased risk of adverse health events in the population is to collect this single measure. However, the other measures do add additional information. If all four HRQOL-4 items are collected, these items could be used to create a better model of mortality risk which could inform targeting of future preventive intervention efforts. For example, if these items were part of a clinical screening form or a routine wellness questionnaire, persons with worsening over repeated measures of GSRH, PUDs, or ALDs could be flagged for additional clinician or caseworker review to better understand their health state and higher risk of mortality. The inconsistent and largely non-significant relationship for mental health suggests that less emphasis should be put on the MUD measure when monitoring for population health risk. However, it is still an important outcome capturing quality of life associated with morbidity.

This research has limitations. The large sample of Medicare recipients is an advancement, but is not representative of either the older adult population or the entire Medicare population. Our sample restrictions change the composition slightly relative to the HOS sampling frame, but we have no reason to believe that HRQOL–mortality relationship is affected by this. Second, time to mortality is measured with error among mail respondents: HOS records the date on which the survey was received, not when it was completed. This short lag should have minimal impact on the analysis. The proportional hazards assumption was not met for the maximum follow-up time or for GSRH in the short-term model, so final conclusions about the HRs for HRQOL should be drawn from Table 4, which includes time interactions. Finally, the back pain relationship in the multivariate models is puzzling. Back pain is correlated with some of the other controls in the multivariate model. Our suspicion is that after including the

12 chronic diseases, GSRH, and HDs, the residual information in back pain in the multivariate model is small and may be proxying for attitudinal differences in the sample. Because our main emphasis was on the HRQOL relationship, further exploration of this finding must be left for future research.

Future research may address these limitations in several ways. Population differences are likely to be important and may be explored with different samples and greater collection of HRQOL measures, including the Healthy Days. The relationships may also be different for less healthy Medicare beneficiaries and could alter conclusions about using these results for broader population surveillance. The question of proxy respondents may also be an important area for future research. In conclusion, this study demonstrates strong associations between HRQOL and mortality using the CDC Healthy Days measures in both short- and long-term measurement, with important differences identified over time.

## References

- Albrecht, G. L., & Devlieger, P. J. (1999). The disability paradox: High quality of life against all odds. *Social Science and Medicine*, *48*, 977–988.
- Andresen, E. M., Vahle, V. J., & Lollar, D. (2001). Proxy reliability: Health-related quality of life (HRQOL) measures for people with disability. *Quality of Life Research*, *10*, 609–619.
- Atherly, A., Hebert, P. L., & Maciejewski, M. L. (2005). An analysis of disenrollment from Medicare managed care plans by Medicare beneficiaries with diabetes. *Medical Care*, *43*, 500–506.
- Bowling, A. (1991). *Measuring health: A review of quality of life measurement scales*. Bristol, PA: Open University Press.
- Cavrini, G., Broccoli, S., Puccini, A., & Zoli, M. (2012). EQ-5D as a predictor of mortality and hospitalization in elderly people. *Quality of Life Research*, *21*, 269–280.
- CDC (Centers for Disease Control and Prevention). (2000). *Measuring healthy days*. Atlanta, GA: CDC.
- Cherepanov, D., Palta, M., & Fryback, D. G. (2010). Underlying dimensions of the five health-related quality-of-life measures used in utility assessment: Evidence from the National Health Measurement Study. *Medical Care*, *48*, 718–725.
- Corti, M.-C., Guralnik, J. M., Ferrucci, L., Izmirlian, G., Leveille, S. G., Pahor, M., et al. (1999). Evidence for a black-white crossover in all-cause and coronary heart disease mortality in an older population: The North Carolina EPESE. *American Journal of Public Health*, *89*, 308–314.
- Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society, Series B*, *34*, 187–220.
- DeSalvo, K. B., Bloser, N., Reynolds, K., He, J., & Muntner, P. (2006a). Mortality prediction with a single general self-rated health question. A meta-analysis. *Journal of General Internal Medicine*, *21*, 267–275.
- DeSalvo, K. B., Fan, V. S., McDonnell, M. B., & Fihn, S. D. (2005). Predicting mortality and healthcare utilization with a single question. *Health Services Research*, *40*, 1234–1246.
- DeSalvo, K. B., Fisher, W. P., Tran, K., Bloser, N., Merrill, W., & Peabody, J. (2006b). Assessing measurement properties of two

- single-item general health measures. *Quality of Life Research*, 15, 191–201.
- Dominick, K. L., Ahern, F. M., Gold, C. H., & Heller, D. A. (2002). Relationship of health-related quality of life to health care utilization and mortality among older adults. *Aging Clinical and Experimental Research*, 14, 499–508.
- Ellis, B. H., Bannister, W. M., Cox, J. K., Fowler, B. M., Shannon, E. D., Drachman, D., et al. (2003). Utilization of the propensity score method: An exploratory comparison of proxy-completed to self-reported responses in the Medicare Health Outcomes Survey. *Health and Quality of Life Outcomes*, 1, 47. doi:10.1186/1477-7525-1-47.
- Fayers, P., & Machin, D. (2007). *Quality of life: The assessment, analysis, and interpretation of patient-reported outcomes*. New York: Wiley.
- Franks, P., Gold, M. R., & Fiscella, K. (2003). Sociodemographics, self-rated health, and mortality in the U.S. *Social Science and Medicine*, 56, 2505–2514.
- Fryback, D. G., Dunham, N. C., Palta, M., Hanmer, J., Buechner, J., Cherepanov, D., et al. (2007). U.S. norms for six generic health-related quality-of-life indexes from the National Health Measurement Study. *Medical Care*, 45, 1162–1170.
- Gold, M., Phelps, D., Jacobson, G., & Neuman, T. (2010). Medicare Advantage 2010 data spotlight. Plan enrollment patterns and trends. Kaiser Family Foundation Medicare Policy Project, publication 8080. Retrieved from <http://www.kff.org/medicare/upload/8080.pdf>.
- Hanmer, J., Lawrence, W. F., Anderson, J. P., Kaplan, R. M., & Fryback, D. G. (2006). Report of nationally representative values for the noninstitutionalized US adult population for 7 health-related quality-of-life scores. *Medical Decision Making*, 26, 391–400.
- Heinze, G., & Schemper, M. (2003). Comparing the importance of prognostic factors in Cox and logistic regression using SAS. *Computer Methods and Programs in Biomedicine*, 71, 155–163.
- Johnson, N. E. (2000). The racial crossover in comorbidity, disability, and mortality. *Demography*, 37, 267–283.
- Korten, A. E., Jorm, A. F., Jiao, Z., Letenneur, L., Jacomb, P. A., Henderson, A. S., et al. (1999). Health, cognitive, and psychosocial factors as predictors of mortality in an elderly community sample. *Journal of Epidemiology and Community Health*, 53, 83–88.
- Lied, T. R., Sheingold, S. H., Landon, B. E., Shaul, J. A., & Cleary, P. D. (2003). Beneficiary reported experience and voluntary disenrollment in Medicare managed care. *Health Care Financing Review*, 25, 55–66.
- Mello, M. M., Stearns, S. C., Norton, E. C., & Ricketts, T. C., 3rd. (2003). Understanding biased selection in Medicare HMOs. *Health Services Research*, 38, 961–992.
- Moriarty, D. G., Kobau, R., Zack, M. M., & Zahran, H. S. (2005). Tracking Healthy Days—A window on the health of older adults. *Preventing Chronic Disease*, 2, 1–8. Retrieved from [http://www.cdc.gov/pcd/issues/2005/jul/05\\_0023.htm](http://www.cdc.gov/pcd/issues/2005/jul/05_0023.htm).
- Moriarty, D. G., Zack, M. M., & Kobau, R. (2003). The Centers for Disease Control and Prevention's Healthy Days measures—Population tracking of perceived physical and mental health over time. *Health and Quality of Life Outcomes*, 1, 37. doi:10.1186/1477-7525-1-37.
- Netuveli, G., & Blane, D. (2008). Quality of life in older ages. *British Medical Bulletin*, 85, 113–126.
- Pratt, L. A. (2009). Serious psychological distress, as measured by the K6, and mortality. *Annals of Epidemiology*, 19, 202–209.
- Quinten, C., Coens, C., Mauer, M., Comte, S., Sprangers, M. A., Cleeland, C., et al. (2009). Baseline quality of life as a prognostic indicator of survival: A meta-analysis of individual patient data from EORTC clinical trials. *Lancet Oncology*, 10, 865–871.
- Sargent-Cox, K. A., Anstey, K. J., & Luszcz, M. A. (2010). The choice of self-rated health measures matter when predicting mortality: Evidence from 10 years follow-up of the Australian longitudinal study of ageing. *BMC Geriatrics*, 10, 18. doi:10.1186/1471-2318-10-18.
- Schemper, M. (1993). The relative importance of prognostic factors in studies of survival. *Statistics in Medicine*, 12, 2377–2382.
- Schemper, M., & Stare, J. (1996). Explained variation in survival analysis. *Statistics in Medicine*, 15, 1999–2012.
- Sondik, E. J., Huang, D. T., Klein, R. J., & Satcher, D. (2010). Progress toward the Healthy People 2010 goals and objectives. *Annual Review of Public Health*, 31, 271–281.
- Spilker, B. (Ed.). (1995). *Quality of life and pharmacoeconomics in clinical trials* (2nd ed.). Philadelphia, PA: Lippincott, Williams, and Wilkins.
- Stephens, A., & Wardle, J. (2011). Positive affect measured using ecological momentary assessment and survival in older men and women. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 18244–18248.
- Tsai, S. Y., Chi, L. Y., Lee, C. H., & Chou, P. (2007). Health-related quality of life as a predictor of mortality among community-dwelling older persons. *European Journal of Epidemiology*, 22, 19–26.
- Vogt, T., Pope, C., Mullooly, J., & Hollis, J. (1994). Mental health status as a predictor of morbidity and mortality: A 15-year follow-up of members of a health maintenance organization. *American Journal of Public Health*, 84, 227–231.