BACKGROUND: The Health and Retirement Study is a national sample of Americans older than 50 years and their spouses. The present study evaluated cross-sectional and longitudinal data from January 2000 through December 2006. The objective of the study was to evaluate the roles of spouse-rated vs self-rated health as predictors of all-cause mortality among adults older than 50 years.

METHODS: A total of 673 dyads of married couples were randomly selected to participate in a Health and Retirement Study module examining spouse-rated health. For each couple, one member was asked to rate his or her overall health status, and his or her spouse was asked to report the partner's overall health status. Mortality data were available through 2006.

RESULTS: Our findings demonstrate that spouse-rated health (area under the curve, 0.75) is as strong a predictor of mortality as self-rated health (area under the curve, 0.73) ($\chi^2=0.36, P=0.54$). Combining spouse-rated and self-rated health predicts mortality better than using self-rated health alone (area under the curve, 0.77) ($\chi^2=6.72, P=0.009$).

CONCLUSIONS: Spouse ratings of health are at least as strongly predictive of mortality as self-rated health. This suggests that, when self-rated health is elicited as a prognostic indicator, spouse ratings can be used when self-ratings are unavailable. Both measures together may be more informative than either measure alone.

Arch Intern Med. 2009;169(22):2156-2161
The Health and Retirement Study (HRS) is a nationally representative sample of individuals older than 50 years living in the United States. The HRS is sponsored by the National Institute on Aging, Bethesda, Maryland, and is conducted by the University of Michigan, Ann Arbor. The study was reviewed and approved by the University of Michigan’s Health Sciences Institutional Review Board. Participants take part in a biennial interview that covers a range of topics, including income, wealth, work, retirement, health, health care use, and other factors. Most interviews are conducted over the telephone.

The present study evaluated cross-sectional and longitudinal data from January 2000 through December 2006. Baseline data were collected in 2000. Overall, 19,580 individuals responded to the 2000 HRS questionnaire. In addition to the core interview, each wave of the HRS includes additional modules on selected topics that are administered to randomly selected participants. Randomization is computerized and is conducted by the University of Michigan. The potential analytic sample for this study was 747 married participants who were randomly selected to participate in a module designed to compare respondent and spouse responses to questions about respondents’ health status. Of 747 proxies, 66 were excluded because the spouse was expected to be unable to provide data on his or her own health status based on low cognitive performance demonstrated in past waves or known health problems, and 8 were excluded because the spouse was unable to provide data on his or her own health status at the current wave. Therefore, baseline data were available on 673 respondent-spouse dyads. These pairs are representative of all coupled respondents in the 2000 HRS sample. Follow-up mortality data were available through 2006.

OUTCOME VARIABLE

Mortality status (alive or dead) was available through 2006 based on HRS tracking efforts. Respondents were classified by HRS into 1 of the following 5 categories: (1) alive in 2006, (2) presumed alive as of 2006, (3) death reported in 2006, (4) death reported in past waves, or (5) vital status unknown. In the present study, the first 2 categories were collapsed as alive, the next 2 categories were collapsed as dead, and the fifth category was classified as missing (1.1% of the sample).

PREDICTORS

Spouse-Rated Health

Proxies were asked to rate the health status of their spouse. The following 5-point scale was used: 5 (excellent), 4 (very good), 3 (good), 2 (fair), and 1 (poor).

Self-rated Health

All participants were asked to rate their own health status. The same 5-point scale was used.

STATISTICAL ANALYSIS

We performed chi² analyses to compare spouses and respondents on various categorical demographic and clinical characteristics. t Tests for dependent samples were performed to compare the 2 parties on continuous variables. We also used the Wilcoxon signed rank test to compare spouses’ and respondents’ reports of health status. We then evaluated the degree of concordance between spouse-rated health (of the respondent) and self-rated health using chi² analysis and the weighted κ statistic via a linear set of weights (eg, 1.0, 0.75, 0.50, 0.25, and 0). The κ statistic examines the degree of agreement beyond what would occur by chance. A κ statistic of 0 indicates that the level of agreement is no more than would be expected by chance alone, while a κ statistic of 1.0 indicates perfect agreement. Next, we identified the unique association of the various covariates with spouse-rated health.

COVARIATES

Medical Status

The number of chronic medical conditions among 6 common conditions (eg, cancer, diabetes mellitus, and stroke) was gathered based on self-report. We dichotomized subjects as those with 0 to 1 medical conditions vs multiple conditions.

Functional Impairment

Participants were asked to indicate the presence or absence of impairment in 9 activities (eg, difficulties in stooping and difficulties in sitting). Participants were then asked whether they provide assistance with activities of daily living or instrumental activities of daily living to their spouse.

Cognitive Functioning

We used the HRS Cognitive Scale, a test of overall cognitive functioning that includes subtests of immediate and delayed word recall, subtraction, and backward count. The subtests were modeled after the Mini-Mental State Examination, a standard geriatric dementia screen. Because scores are highly correlated with each other, we used a composite score ranging from 0 to 26, with 26 representing perfect performance.

Health Behaviors

Participants were asked whether they participate in any vigorous physical activity, smoke cigarettes, or drink alcohol. Answers were recorded as yes or no.

Sociodemographic Data

Sociodemographic data were based on self-report. Data recorded included sex, age (<65, 65-74, or ≥75 years), education status (0-12 or ≥13 years), and race/ethnicity (white, black, Latino, or other).
over and above their associations with self-rated health. We first ensured that all variables complied with the proportional odds assumption and then performed a series of proportional odds regression analyses with spouse-rated health as an outcome variable, each sociodemographic and clinical correlate as a potential predictor, and self-rated health as a control variable. We then performed logistic regression analyses comparing spouse-rated health with self-rated health as a potential predictor, and self-rated health as a control variable. Only 272 dyads completed the module.

### RESULTS

Overall, 673 dyads participated in the study. Most spouses (54.0%) were female. Most spouses (87.7%) and respondents (86.4%) were white (Table 1). Although differences between respondents and spouses were statistically significant for some characteristics, the magnitude of the difference was small in most cases.

As summarized in Table 2, the degree of concordance between spouse-rated health and self-rated health was moderate (weighted κ = 0.48, P < .001). This suggests that the level of agreement between respondents and spouses was at about the midpoint between no better than chance and perfect. The Wilcoxon signed rank test showed that spouses tended to rate respondents’ health slightly worse than respondents did (mean [SD] score, 3.22 [0.05] by spouses and 3.37 [0.05] by respondents; \( z = 4.3, P < .001 \)).

### CORRELATES OF SPOUSE-RATED HEALTH AFTER CONTROLLING FOR SELF-RATED HEALTH

Using \( P < .01 \) as the level of statistical significance to account for the large sample size and the multiple comparisons,

all variables met the proportional odds assumption. In a series of proportional odds regression analyses, we found several respondent-level and spouse-level variables to be associated with spouse-rated health even after controlling for self-rated health (Table 3). Respondents of younger age, higher education, fewer medical conditions, higher cognitive functioning, and less functional impairment were more likely to be rated by their spouse as enjoying better health, as were those who engaged in physical activity. In addition, nondepressed spouses and spouses with better cognitive functioning were more likely to rate their partner’s health as better independent of their partner’s self-rated health. Hence, these various demographic and clinical characteristics of both respondent and spouse contribute to spousal perception of respondents’ health even after respondents’ self-rated health is considered.

### SPOUSE-RATED HEALTH VS SELF-RATED HEALTH AS PREDICTORS OF MORTALITY

Overall, 94 of 673 respondents (12.3%) died during the 6-year period. As summarized in Table 4, better self-rated health is associated with lower mortality risk (\( F_{4,47} = 12.5, P < .001 \)) based on logistic regression analysis. Similar results were obtained for spouse-rated health (\( F_{4,47} = 18.0, P < .001 \)). Both spouse-rated and self-rated health remained significant predictors of mortality even after adjusting for various demographic and clinical characteristics (eg, respondents’ age, sex, education status, medical status, functional impairment, depression, cognitive status, health behaviors, and cognitive functioning, as well as spouses’ depression, cognitive status, and caregiving status).

Areas under the curve indicated moderate predictive ability for both spouse-rated and self-rated health. Spouse-rated health (area under the curve, 0.75) was not signifi-
significantly better than self-rated health (area under the curve, 0.73) ($\chi^2=0.36$, $P=.54$). Combining spouse-rated and self-rated health provided the best predictive value of mortality (area under the curve, 0.77), which was significantly better than self-rated health alone ($\chi^2=6.72$)

### Table 2. Self-rated Health vs Spouse-Rated Health

<table>
<thead>
<tr>
<th>Variable</th>
<th>Poor</th>
<th>Fair</th>
<th>Good</th>
<th>Very Good</th>
<th>Excellent</th>
<th>Total, No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-rated health</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td>27</td>
<td>9</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>40 (4.9)</td>
</tr>
<tr>
<td>Fair</td>
<td>22</td>
<td>51</td>
<td>30</td>
<td>3</td>
<td>2</td>
<td>108 (15.5)</td>
</tr>
<tr>
<td>Good</td>
<td>6</td>
<td>44</td>
<td>96</td>
<td>45</td>
<td>13</td>
<td>204 (29.3)</td>
</tr>
<tr>
<td>Very good</td>
<td>1</td>
<td>14</td>
<td>71</td>
<td>105</td>
<td>28</td>
<td>219 (35.5)</td>
</tr>
<tr>
<td>Excellent</td>
<td>1</td>
<td>0</td>
<td>20</td>
<td>34</td>
<td>34</td>
<td>89 (15.5)</td>
</tr>
<tr>
<td>Total, No. (%)</td>
<td>57 (7.0)</td>
<td>118 (17.2)</td>
<td>221 (32.4)</td>
<td>187 (20.0)</td>
<td>77 (11.2)</td>
<td>660</td>
</tr>
</tbody>
</table>

*Weighted $\kappa$ was calculated to evaluate the degree of concordance between spouse and respondent. To identify differences between spouse and respondent, a weighted $\chi^2$ test was performed. Weighted $\kappa = 0.48$, $P < .001$; $\chi^2 = 461.8$, $P < .001$. Percentages do not total 100% due to rounding.

### Table 3. Correlates of Spouse-Rated Health After Adjustment for Self-rated Health

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Odds Ratio (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td></td>
</tr>
<tr>
<td>&lt;65</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>65-74</td>
<td>0.63 (0.44-0.92)</td>
</tr>
<tr>
<td>≥75</td>
<td>0.50 (0.31-0.79)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>Female</td>
<td>1.11 (0.79-1.54)</td>
</tr>
<tr>
<td>Education, y</td>
<td></td>
</tr>
<tr>
<td>0-12</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>≥13</td>
<td>1.67 (1.17-2.37)</td>
</tr>
<tr>
<td>Medical conditions</td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>≥2</td>
<td>0.53 (0.38-0.74)</td>
</tr>
<tr>
<td>Functional impairment</td>
<td></td>
</tr>
<tr>
<td>Not depressed</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>Depressed</td>
<td>0.68 (0.44-1.06)</td>
</tr>
<tr>
<td>Vigorous exercise</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>Yes</td>
<td>1.47 (1.07-2.03)</td>
</tr>
<tr>
<td>Smoking cigarettes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>Yes</td>
<td>0.79 (0.52-1.20)</td>
</tr>
<tr>
<td>Drinking alcohol</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>Yes</td>
<td>0.79 (0.52-1.20)</td>
</tr>
<tr>
<td>Cognitive status</td>
<td></td>
</tr>
<tr>
<td>1.02 (1.02-1.08)</td>
<td></td>
</tr>
<tr>
<td>Spouse’s depression</td>
<td></td>
</tr>
<tr>
<td>Not depressed</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>Depressed</td>
<td>0.43 (0.28-0.64)</td>
</tr>
<tr>
<td>Spouse’s cognitive status</td>
<td></td>
</tr>
<tr>
<td>1.06 (1.02-1.10)</td>
<td></td>
</tr>
<tr>
<td>Providing assistance with ADLs and IADLs to respondent</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>Yes</td>
<td>0.69 (0.32-1.46)</td>
</tr>
</tbody>
</table>

### Table 4. Self-rated Health vs Spouse-Rated Health as Predictors of Mortality

<table>
<thead>
<tr>
<th>Variable</th>
<th>Overall Incidence, No. (%)</th>
<th>Odds Ratio (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-rated health</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td>16 (36.1)</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>Fair</td>
<td>34 (29.4)</td>
<td>0.72 (0.30-1.76)</td>
</tr>
<tr>
<td>Good</td>
<td>24 (11.0)</td>
<td>0.22 (0.09-0.48)</td>
</tr>
<tr>
<td>Very good</td>
<td>14 (5.3)</td>
<td>0.09 (0.03-0.25)</td>
</tr>
<tr>
<td>Excellent</td>
<td>3 (2.3)</td>
<td>0.04 (0.009-0.180)</td>
</tr>
<tr>
<td>Spouse-rated health</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td>28 (45.6)</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>Fair</td>
<td>30 (25.0)</td>
<td>0.39 (0.19-0.79)</td>
</tr>
<tr>
<td>Good</td>
<td>23 (9.7)</td>
<td>0.12 (0.06-0.23)</td>
</tr>
<tr>
<td>Very good</td>
<td>9 (3.6)</td>
<td>0.04 (0.01-0.12)</td>
</tr>
<tr>
<td>Excellent</td>
<td>4 (3.6)</td>
<td>0.04 (0.01-0.16)</td>
</tr>
</tbody>
</table>

Abbreviations: ADLs, activities of daily living; IADLs, instrumental ADLs.

Several studies to date have evaluated the role of self-rated health as a predictor of all-cause mortality, and its predictive ability has been demonstrated in various epidemiologic studies. To our knowledge, this is the first study to show that among older adults spouse-rated health is as predictive of mortality as self-rated health, despite the fact that the 2 are not synonymous. Our findings indicate only moderate levels of concordance between spouse-rated health and self-rated health. Furthermore, the study demonstrates that spouse-rated health is correlated with other respondent-level and spouse-level variables in addition to self-rated health; spouse-rated health captures respondents’ medical status, demographic characteristics, cognitive status, health behav-

©2009 American Medical Association. All rights reserved.
self-rated health vs spouse-rated health.

vs combined spouse-rated and self-rated health.

spouse may have heard his or her partner providing an-

the questions have similar response choices. Although a

of the HRS interview takes a fairly long time, and many of

is unlikely that it had a major effect. The administration

sure how this might have affected the results, we think it

were in place, and it is possible that spouses were present

dents. Therefore, results cannot be generalized to cogni-

representative of individuals 50 years or younger. Second, the

vidual-rated and proxy-rated variables.8 Hence, the present

results may not be generalizable to other types of proxies.

Furthermore, the HRS is a representative sample of indi-

mortality. For this analysis, self-rated and spouse-rated health were recoded as

1 (excellent), 2 (very good), 3 (good), 4 (fair), or 5 (poor). χ² = 36, P = .04 for

self-rated health vs spouse-rated health. χ² = 6.72, P = .009 for self-rated health vs

combined spouse-rated and self-rated health. χ² = 2.91, P = .08 for

spouse-rated health vs combined spouse-rated and self-rated health.

iors, and mental status. As demonstrated in past research,31

spouse-rated health also is correlated with proxies’ cognitive

status and mental status, so that more depressed or cogni-

tively impaired proxies tend to report their spouse’s health

as poorer even after controlling for self-rated health.

Similar to the case of self-rated health, it remains un-

clear what accounts for the strong predictive value of spouse-

ated health. Are spouses more attuned to certain biologic

and physiologic processes in their partner that remain oth-

erwise unnoticed? Is it the mental condition of the spouse

that has such an important effect on respondent mortality

prospects? Or is it the relationship between the 2 partners

and the expectations that spouses hold regarding their part-

ner that affect their partner’s mortality prospects? At this

point, the exact mechanism behind the predictive ability

of spouse-rated health remains unclear. Nevertheless, our

study shows that combining spouse-rated health with self-

rated health provides a better prognostic indicator of all-

cause mortality than self-rated health alone.

The present study has several limitations that should be

acknowledged. First, the study was limited to married

couples. Research has shown that the quality of the rela-

tionship between proxy and respondent has a major role in

determining the degree of concordance between respon-

dent-rated and proxy-rated variables.8 Hence, the present

results may not be generalizable to other types of proxies.

Furthermore, the HRS is a representative sample of indi-

viduals older than 50 years; hence, results may not be re-

presentative of individuals 50 years or younger. Second, the

study was limited to cognitively intact proxies and respon-

dents. Therefore, results cannot be generalized to cogni-

tively impaired participants. Third, no blinding measures

were in place, and it is possible that spouses were present

during their partner’s interview. Although we cannot be

sure how this might have affected the results, we think it

is unlikely that it had a major effect. The administration

of the HRS interview takes a fairly long time, and many of

the questions have similar response choices. Although a

spouse may have heard his or her partner providing an-

swers, it is unlikely that he or she would have known which

specific question was being answered. Perhaps most im-

portant, our results emphasize that spouse ratings of health

are often discordant. As a result, any bias resulting from

awareness of partner response would bias our results to-

ward the null, suggesting that our results are even more

conservative. Fourth, the study did not evaluate the pre-

dictive ability of spouse-rated health for purposes other than

mortality. Fifth, we acknowledge that there are no stan-

dards for how to recognize the clinical significance of dif-

ferences in receiver operating characteristic curves. How-

ever, many investigators would view the receiver operating

characteristic curve difference of 0.04 (self-rated health com-

pared with combined spouse-rated and self-rated health) as

clinically meaningful. Consider a situation in which you

have pairs of patients, one of whom survived longer than

the other. Using self-rated health only, you would cor-

rectly identify the longer surviving patient 73% of the time.

Using the combined spouse-rated and self-rated health

model, you would correctly identify the longer surviving

patient 77% of the time. We believe that this is meaning-

ful and that most would choose to incorporate both spouse-

rated and self-rated health based on this information.

Nevertheless, this is the first study to date to evaluate

the role of spouse-rated health as a predictor of mortal-

ity. Our findings demonstrate that spouse-rated health

is at least as strong a predictor of mortality as self-rated

health, although the 2 measure different entities: spouse-

ated health not only correlates with self-rated health but

also captures the sociodemographic and medical status

of respondents, as well as spouses’ own cognitive and men-

tal status. Spouse-rated health can be used as a predic-

tor of mortality when self-rated health is unavailable or

as an additional source of data that complements self-

rated health. Health care practitioners working with older

adults should attempt to obtain not only patients’ self-

report of their health status but also their spouses’ self-

report whenever available, as the combination of spouse-

rated and self-rated health provides a more accurate

estimate of respondents’ mortality risk.

Accepted for Publication: August 18, 2009.

Correspondence: Liat Ayalon, PhD, The Louis and Gabi

Weisfeld School of Social Work, Bar-Ilan University, Ra-

mat Gan, Israel 52900 (ayalonl@mail.biu.ac.il).

Author Contributions: Study concept and design: Aya-

lon and Covinsky. Analysis and interpretation of data: Aya-

lon and Covinsky. Drafting of the manuscript: Ayalon. Criti-

cal revision of the manuscript for important intellectual

content: Ayalon and Covinsky. Statistical analysis: Aya-

lon and Covinsky. Obtained funding: Covinsky. Adminis-

trative, technical, or material support: Covinsky.

Financial Disclosure: None reported.

Funding/Support: This study was supported by grants

5K24 AG029812 and 5R01 AG023626 from the Na-

tional Institute on Aging (Dr Covinsky).

REFERENCES

1. Strawbridge WJ, Wallhagen MI. Self-rated health and mortality over three de-


21(3):402-416.


