Are Gender Differences in the Relationship Between Self-Rated Health and Mortality Enduring? Results From Three Birth Cohorts in Melton Mowbray, United Kingdom

Antony Arthur, PhD, Carol Jagger, PhD, Michael Clarke, FRCP, Nicola Spiers, PhD
Are Gender Differences in the Relationship Between Self-Rated Health and Mortality Enduring? Results From Three Birth Cohorts in Melton Mowbray, United Kingdom

Nicola Spiers, PhD,1 Carol Jagger, PhD,1 Michael Clarke, FRCP,1 and Antony Arthur, PhD2

Purpose: The purpose of this study was to assess whether there is an enduring gender difference in the ability of self-rated health to predict mortality and investigate whether self-reported physical health problems account for this difference. Design and Methods: Cox models for 4-year survival were fitted to data from successive cohorts aged 75–81 years registered with a primary care practice in the U.K. Midlands surveyed in 1981, 1988, and 1993–1995. Results: Self-rated health was consistently a stronger predictor in men (hazard ratio [HR] = 2.7; 95% confidence interval [CI] = 2.1–3.5) than it was in women (HR = 1.9; 95% CI = 1.5–2.4). Women surveyed in 1993–1995 were more likely than men to report problems that were disabling but not life-threatening, whereas men were more likely to report potentially life-threatening problems. However, these differences did not explain the association of self-rated health with mortality. More than half of those who reported a potentially life-threatening problem said that their health was good. Implications: Self-rated health is more strongly associated with mortality in men, but this is unlikely to be explained by differences in the nature of their physical health problems.

Key Words: Aged, Self-report, Self-perceived health, Cox regression

Gender differences in the ability of self-rated health to predict mortality are of interest because they throw light on the knowledge and experience from which older men and women draw when answering this deceptively simple question. Studies relating self-rated health to mortality have been reviewed by Idler and Benyamini (Benyamini & Idler, 1999; Idler & Benyamini, 1997). In 17 studies where measures were reported separately by gender, the association was stronger in men in 11 studies. Of six studies that reported greater predictive validity for women, the two that had the largest differences included people of working age (Miilunpalo, Vuori, Oja, Pasanen, & Urponen, 1997; Strawbridge & Wallhagen, 1999).

Women suffer more chronic problems in old age (e.g., Verbrugge, 1986; Waldron, 1983) and have greater prevalence of disability (Manton, 1988), whereas men have a higher incidence of life-threatening problems such as cancer and chronic heart disease. Benyamini, Leventhal, and Leventhal (2000) distinguish “serious medical conditions” that are associated with both less than good self-rated health and mortality from nonsevere illnesses that have little relation to mortality in those older than 65 years, but are associated with lower self-rated health. Dimensions of health that are unrelated to mortality have been thought to be more important constituents of self-rated health in women than in men (Jylhä, Guralnik, Ferrucci, Jokela, & Heikkinen, 1998). We investigate the hypothesis that among those in less...
than good self-rated health, there is higher prevalence in women of disabling illness that threatens quality rather than quantity of life, resulting in a weaker association of self-rated health with mortality in women.


Methods

The market town and rural environs of Melton Mowbray, Leicestershire, United Kingdom, are served by a single general medical practice with a practice list size of around 33,000. The first survey population was drawn on December 31, 1980, comprising all those aged 75 and older who were registered with the practice. Over the following 6 months, trained fieldworkers interviewed these participants. Seven years later, the survey was repeated with the population that was aged 75 years and older on December 31, 1987. The current analysis includes people aged 75–81 years on the dates the population was drawn. Descriptions of the methods of both surveys have been published elsewhere (Clarke, Odell, & Jagger, 1984; Jagger, Clarke, & Clarke, 1991).

Since 1990, all those aged 75 years and older and registered with the practice have been offered a routine health assessment approximately every 18 months, carried out by a nurse in the patient’s home. Full details of the assessments can be found elsewhere (Jagger, Clarke, O'Shea, & Gannon, 1996). The third wave of these assessments took place between 1993 and 1995 with the earliest Wave 3 assessment in November 1993 and the latest in November 1995. The practice catchment area is divided into geographical zones that were sampled in succession over the assessment period. We included in our analysis those in the sample who were alive and aged 75–81 on December 31, 1994. The three surveys provided successive birth cohorts (born 1899–1905, 1906–1912, and 1913–1919) aged 75–81 years when the survey took place.

Measures

We obtained fact and date of death from the National Health Service Central Registry. We assessed self-rated health by answers to the question “For your age, would you say that your health in general was good, fair or poor?” with responses dichotomized for analysis into good and fair or poor. Baseline sociodemographic variables included sex and whether the person was living alone. We defined activity restrictions as having difficulty performing alone, using help or aids, or not doing each of six activities of daily living (ADLs): (a) mobility around the home, (b) transfer to and from chair, (c) transfer to and from bed, (d) feeding, (e) dressing, and (f) bathing. We categorized physical function into good (able to perform all of the six activities without help from another person or appliance), moderate (able to perform all of the six activities without help from another person, but with help from an appliance in at least one activity), or poor (having difficulty, needing help from another person, or being unable to do at least one of the six activities).

We assessed cognitive function by the Information/Orientation (I/O) subtest of the Clifton Assessment Procedures for the Elderly (Pattie & Gilleard, 1979), a 12-item screen for cognitive impairment. Scores of 8 and under have been shown to correspond well to moderate or severe dementia (Jagger, Clarke, & Anderson, 1992). We restricted the range of health covariates for which we adjusted to those available at the time of the first survey, but also included urinary incontinence and difficulty hearing (hearing difficulty or impaired hearing, even with aid).

For the health assessments only, information was also available on 18 troublesome health problems, for example, weight loss, problems with chest or breathing, and cramps on walking. For each troublesome health problem in turn, we asked participants whether they had a problem, with further description being given by the nurse to aid understanding where requested.

Statistical Methods

First, we fitted Cox Proportional Hazards models for 4-year survival separately for the 1981 and 1988 and 1993–1995 cohorts. We measured time from date of interview to date of death, censored at 4 years from interview. We estimated gender differences in the self-rated health–mortality association by including an interaction term between gender and self-rated health. We repeated this process with adjustment for the following known predictors of mortality: (a) ADL dependency, (b) cognitive impairment, (c) urinary incontinence, (d) difficulty hearing, and (e) difficulty seeing. We checked the proportional hazards assumption by including time-varying covariates for all main effects in the model. We then tested for all two-way interactions between main effects. For significant interactions, we recorded the effect of including the interaction term upon the gender-specific risk ratios.

Second, we included responses from all cohorts in a single Cox model for 4-year survival. To account for improvements in mortality in the newer cohorts, we stratified the baseline hazard by cohort. This model
makes minimal assumptions about changes in the age-specific hazard function across cohorts. We used it to assess evidence for a three-way interaction between cohort, self-rated health, and gender, and to derive gender-specific estimates of the self-rated health–mortality relationship aggregated across cohorts.

To investigate the role of self-reported physical health problems in accounting for the gender difference in the self-rated health–mortality relationship, we first ordered the health problems by strength of association with mortality, by using a Cox regression model for all participants in 1993–1995. We used two thresholds to classify problems as disabling or potentially life-threatening according to strength of association with mortality. We then classified participants as having no self-reported problem, having at least one disabling problem but no potentially life-threatening problem, or having a potentially life-threatening problem. Within these subgroups, we estimated the strength of the self-rated health–mortality relationship by gender.

Results

In the cohort drawn on December 31, 1980, we identified 1,329 people aged 75 years and older on the age–sex register. Of these, 50 (4%) died before interview, 11 (1%) could not be contacted, and 65 (5%) refused. We did not ask those interviewed in institutions to rate their health and excluded them from the analysis (9%, n = 139). We excluded a further 45 (2%) because they were aged less than 75 years on December 31, 1994.

We interviewed the following numbers at home who were aged 75–81 years at the date the cohort was defined: 823 in the 1981 cross-section, 995 in the 1988 cross-section, and 803 in the 1993–1995 health assessments. The cohorts we included in the analysis and the losses because of missing data are summarised in Table 1. For 19 individuals where values were missing for one only of the six ADLs, we calculated a reweighted score, based upon the five nonmissing responses.

In each cohort, women with good self-rated health had the lowest mortality followed by men in good self-rated health and women in poor or fair self-rated health, with men in poor or fair self-rated health facing the greatest risk. Hazard ratios from the Cox models for 4-year survival are given in Table 2. The association of mortality with poor or fair self-rated health was stronger for men than for women in each cohort, although the Gender × Self-Rated Health interaction was only statistically significant in those born in 1906–1912. The Cohort × Gender × Self-Rated Health interaction is small and statistically insignificant, so a conclusion of an enduring gender difference is consistent with these data.

There is some evidence, albeit inconsistent across the cohorts, that the association between self-rated health and mortality is time varying, with a stronger effect upon early than late mortality. Allowing for a significant time-varying effect of hearing impairment and for five significant two-way interactions between covariates had no substantial effect upon the gender difference. Results are reported for a parsimonious model that assumes that the ratio of mortality for good versus fair or poor health is constant over time.

Prevalence of troublesome physical health problems in the 1993–1995 cohort are shown in Table 3.
The five most prevalent problems were severe joint pain or problems (39.4%), problems with chest or breathing (25.5%), problems with feet or cutting toenails (24.0%), blood pressure problems (19.3%), and very swollen ankles (13.7%). A Cox model was fitted to rank the problems according to strength of association with mortality. The 5 problems most closely associated with mortality, with hazard ratios greater than 1.5, were classified as potentially life threatening, and the remaining 12 problems were classified as disabling. We divided participants into three groups: (a) those who reported none of the 17 specific troublesome problems, (b) those who reported at least one specific disabling problem but no life-threatening problem, and (c) those who reported at least one life-threatening problem (Table 4). As expected, prevalence of disabling problems was greater among women (82.1% versus 72.9% reported at least one disabling problem), and prevalence of potentially life-threatening problems was greater among men (42.7% versus 32.1% reported at least one life-threatening problem).

We fitted a Cox model for the 1993–1995 cohort, including interaction terms to estimate the relative risk of mortality for those in less than good self-rated health separately by gender and type of health problem. Although type of problem was closely related to mortality by construction, including this variable in the model had little impact on the relative risk associated with less than good self-rated health by gender (Table 2). The gender difference in the ability of self-rated health to predict mortality was largely attributable to those who reported disabling but not life-threatening problems (Table 4).

If type of problem accounts for the gender difference in the ability of self-rated health to predict mortality, then we would expect to see a reduced gender gap within the categories for this variable. In fact, among those with disabling problems only, the gender gap remained large, with self-rated health being strongly related to mortality among men, but a much weaker association for women. Among those with a life-threatening problem, self-rated health was only weakly associated with mortality for both genders. More than half (55.5%) of those who reported a potentially life-threatening problem also said that their health was good. Because the threshold for a life-threatening problem was chosen arbitrarily, the analysis was repeated, with the eight problems most closely associated with mortality classified as life threatening. The conclusions remained unchanged.

### Discussion

These results show an enduring gender difference over successive birth cohorts in the ability of self-rated health to predict mortality, with the association being stronger for men than women, both in raw comparisons, and when the effect of other health measures is allowed for. The gender difference is of consistent direction across separate cohorts aged 75–81 years drawn from the same geographic population in 1981, 1988, and 1993–1995. Although the sampling frame is similar to England and Wales in age–gender and social class distribution, some caution is in order in generalizing these findings, as the study is limited to those registered with a single medical practice.

There are some further limitations arising partly from the age of the study. The range of health covariates allowed for is quite restricted, and there is a change of study design to take into account, with the 1993–1995 cohort being drawn from a continuous program of health assessments rather than a
cross-sectional survey. This is reflected in the lower figures for deaths and noncontacts, and higher figure for refusals in the 1993–1995 cohort. In this population, older people who decline a health assessment have been shown to have similar functional status, but lower consultation rates, better self-rated health and higher morale, compared with those taking part (Jagger et al., 1996). There may be some bias in the results for 1993–1995 because of lower take-up among those with good self-rated health, but given that response rates are similar for men and women, serious bias is unlikely.

The findings add to the balance of studies of those aged 75 years and older pointing to a stronger association of self-rated health with survival for men, but factors accounting for this gender difference remain unclear. We hypothesized that self-rated health was less strongly associated with mortality in older women, because they suffer disproportionately from problems that threaten quality rather than quantity of life. However, the data on self-reported conditions in 1993–1995 provide scant support for this hypothesis.

Type of problem reported does not account for the gender difference in the self-rated health–mortality association, as the gender difference persists among the majority who have disabling but no life-threatening problems. This is unlikely to be explained by men’s underreporting of life-threatening problems, as a greater prevalence of life-threatening problems is observed in men.

Alternative explanations for the gender difference lie at the boundaries of the list of self-reported health problems. Although the list is nonstandard, symptoms that indicate the most common serious physical conditions in this age group are covered. A possible exception is genitourinary symptoms, but incontinence of urine is allowed for in the analysis. However, severity of problem and diagnosis are also omitted. Benyamini, Leventhal, and Leventhal (2000)

### Table 3. Troublesome Problems by Gender and Association With Mortality: 1993–1995 Health Assessment

<table>
<thead>
<tr>
<th>Troublesome Problem</th>
<th>Rank</th>
<th>Hazard Ratio</th>
<th>Women, % (n = 474)</th>
<th>Men, % (n = 329)</th>
<th>Relative Risk (Men/Women)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight loss</td>
<td>1</td>
<td>4.0***</td>
<td>3</td>
<td>5</td>
<td>1.7</td>
</tr>
<tr>
<td>Weakness/numbness of limbs/stroke</td>
<td>2</td>
<td>3.0***</td>
<td>4</td>
<td>6</td>
<td>1.5</td>
</tr>
<tr>
<td>Passing/coughing/vomiting blood</td>
<td>3</td>
<td>2.2</td>
<td>0.4</td>
<td>0.9</td>
<td>2.3</td>
</tr>
<tr>
<td>Diabetes</td>
<td>4</td>
<td>2.1**</td>
<td>6</td>
<td>8</td>
<td>1.3</td>
</tr>
<tr>
<td>Chest/breathing</td>
<td>5</td>
<td>1.7***</td>
<td>23</td>
<td>29</td>
<td>1.3</td>
</tr>
<tr>
<td>Serious constipation</td>
<td>6</td>
<td>1.5</td>
<td>7</td>
<td>3</td>
<td>0.4*</td>
</tr>
<tr>
<td>Very swollen ankles</td>
<td>7</td>
<td>1.3</td>
<td>18</td>
<td>8</td>
<td>0.4**</td>
</tr>
<tr>
<td>Pain on exertion</td>
<td>8</td>
<td>1.3</td>
<td>12</td>
<td>12</td>
<td>1.0</td>
</tr>
<tr>
<td>Very unsteady/number of falls</td>
<td>9</td>
<td>1.1</td>
<td>7</td>
<td>5</td>
<td>0.7</td>
</tr>
<tr>
<td>Other problem</td>
<td>10</td>
<td>1.1</td>
<td>24</td>
<td>19</td>
<td>0.8</td>
</tr>
<tr>
<td>Severe joint problem</td>
<td>11</td>
<td>1.0</td>
<td>45</td>
<td>31</td>
<td>0.7**</td>
</tr>
<tr>
<td>Allergies</td>
<td>12</td>
<td>0.94</td>
<td>7</td>
<td>5</td>
<td>0.7</td>
</tr>
<tr>
<td>Problem with feet/cutting toenails</td>
<td>13</td>
<td>0.94</td>
<td>26</td>
<td>21</td>
<td>0.8</td>
</tr>
<tr>
<td>Troublesome skin problems</td>
<td>14</td>
<td>0.86</td>
<td>9</td>
<td>14</td>
<td>1.6*</td>
</tr>
<tr>
<td>Diarrhea/loose motions</td>
<td>15</td>
<td>0.86</td>
<td>5</td>
<td>3</td>
<td>0.6</td>
</tr>
<tr>
<td>Blackouts/serious dizzy spells</td>
<td>16</td>
<td>0.83</td>
<td>12</td>
<td>9</td>
<td>0.8</td>
</tr>
<tr>
<td>Cramps on walking</td>
<td>17</td>
<td>0.80</td>
<td>5</td>
<td>8</td>
<td>1.6</td>
</tr>
<tr>
<td>Blood pressure problems</td>
<td>18</td>
<td>0.76</td>
<td>22</td>
<td>15</td>
<td>0.7*</td>
</tr>
</tbody>
</table>

*p < .05; **p < .01; ***p < .001.

### Table 4. Hazard Ratios for Mortality Associated With Less Than Good Self-Rated Health by Gender and Type of Health Problem: 1993–1995 Health Assessment

<table>
<thead>
<tr>
<th>Type of Problems Reported</th>
<th>Women</th>
<th>Men</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n(%)</td>
<td>HR</td>
</tr>
<tr>
<td>None</td>
<td>73(15.4)</td>
<td>2.91</td>
</tr>
<tr>
<td>Disabling problem only</td>
<td>249(52.5)</td>
<td>1.54</td>
</tr>
<tr>
<td>Potentially life-threatening problema</td>
<td>152(32.1)</td>
<td>1.42</td>
</tr>
</tbody>
</table>

Note: HR = hazard ratio; CI = confidence interval.

*aAt least one of: 1. Weight loss; 2. Weakness or numbness of limbs or stroke; 3. Passing, coughing or vomiting blood; 4. Diabetes; 5. Problem with chest or breathing.

*bAdjusted for living alone, activity of daily living dependency, urinary incontinence, cognitive impairment, and hearing impairment.

*p < .05.
categorized diseases as mild or severe and used a scale based on medical history that included measures of severity, but even so they found that allowing for differences in severity of mild and serious diseases had little or no bearing on the gender difference in the self-rated health—mortality relationship. However, the severity assessments were carried out by medical internists on the basis of reported conditions and may have only partially captured more subjective aspects of severity, such as pain and threat to future activities.

A further aspect of health that we do not include in the present analysis is psychosocial functioning. Qualitative research on self-rated health has shown that people may draw on broader aspects of health in addition to physical problems and functioning when rating their health (Jylhä, 1996; Krause & Jay, 1994). Although a measure of restrictiveness and/or inclusiveness of health definition has been developed (Idler, Hudson, & Leventhal, 1999), gender differences in inclusiveness have yet to be explored. However, a measure aggregating fatigue, anxiety, and depression accounted for the gender difference in the self-rated health—mortality association in the Rutgers Ageing and Health Study (Benyamini, Leventhal, & Leventhal, 2000).

The classification of self-reported problems as life threatening or disabling was developed on the present population, although the problems identified indicate cancer, stroke, diabetes, and respiratory problems, reflecting the leading causes of death in those aged 65 and older in the United Kingdom, with the exception of acute myocardial infarction (U.K. National Statistics, 1999). In the absence of replication in other populations, conclusions from the analysis by type of problem in Table 4 must be tentative. Nevertheless, the weakness of the self-rated health—mortality association in the subgroup who reported a life-threatening problem is of interest, given the moderate prevalence of less than good health in this group. Reflecting on qualitative data, Jylhä (1996) argues that “elements incorporated into the concept of health may change quite radically when the individual falls ill” (p. 990). A subgroup of health optimists, who maintain good self-rated health in the face of poor medical ratings has long been recognized (Maddox, 1962). These data emphasize that less than good self-rated health is an imperfect indicator of the presence of threatening symptoms. However, the finding in other studies (Borawski, Kinney, & Kahana, 1996; Chipperfield, 1993) that health optimists have significantly improved survival compared with realists in poor health who rate their health accordingly, is only weakly supported here.

The findings provide further evidence for the hypothesis that self-rated health is a better predictor of survival in men than in women. They suggest that reasons to account for this gender difference are more likely to be found in variation in the definitions that individuals call upon when rating their health, than in differences in the burden of physical morbidity. Although self-rated health remains a useful population health summary, more work is necessary to clarify its interpretation in subgroups suffering from serious ill health.

References

Received November 20, 2001
Accepted April 15, 2002
Decision Editor: Laurence G. Branch, PhD