Social Science & Medicine 187 (2017) 58–66

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Gender and the structure of self-rated health across the adult life span

Anna Zajacova [a, *], Snehalata Huzurbazar [b], Megan Todd [c]

[a] Western University, Canada
[b] University of Wyoming, United States
[c] Columbia University, United States

A R T I C L E I N F O

Article history:
Received 1 February 2017
Received in revised form
21 April 2017
Accepted 15 June 2017
Available online 16 June 2017

Keywords:
Self-rated health
Gender
Validity
Life span
US adults
Bayesian

A B S T R A C T

Despite the widespread use of self-rated health (SRH) in population health studies, the meaning of this holistic health judgment remains an open question. Gender differences in health, an issue of utmost importance in population research and policy, are often measured with SRH; the comparisons could be biased if men and women differ in how they form their health judgment. The aim of this study is to examine whether men and women differ in how health inputs predict their health rating across the adult life span.

We use the 2002–2015 National Health Interview Survey data from US-born respondents aged 25–84. Ordered logistic models of SRH as a function of 24 health measures including medical conditions and symptoms, mental health, functioning, health care utilization, and health behaviors, all interacted with gender, test how the measures influence health ratings and the extent to which these influences differ by gender. Using a Bayesian approach, we then compare how closely a select health measure (K6 score) corresponds to SRH levels among men and women.

We find little systematic gender difference in the structure of SRH: men and women use wide-ranging health-related frames of reference in a similar way when making health judgments, with some exceptions: mid-life and older men weigh physical functioning deficits and negative health behaviors more heavily than women. Women report worse SRH than men on average but this only holds through mid-adulthood and is reversed at older ages; moreover, the female disadvantage disappears when differences in socio-economic and health covariates are considered.

Our findings suggest that the meaning of SRH is similar for women and men. Both groups use a broad range of health-related information in forming their health judgment. This conclusion strengthens the validity of SRH in measuring gender differences in health.

1. Gender and the structure of self-rated health across the adult life span

How do men and women form their self-rated health judgment? Do they differ in what health dimensions matter most? And do any observed patterns hold across the adult life span?

Self-rated health (SRH) has been used extensively to study health trends and inequalities between men and women, as well as across other population groups (Benyamini and Idler, 1999; DeSalvo et al., 2006; Idler and Benyamini, 1997). SRH is a single-question item that asks respondents to rate their own health, typically on a five-point scale ranging from “excellent” to “poor.” Key strengths of SRH are its high reliability (Lundberg and Manderbacka, 1996; Zajacova and Dowd, 2011) and criterion validity. The latter implies that SRH correlates highly with concurrent measures of health (Cousins, 1997; Gold et al., 1996; Jylha et al., 2006) and also predicts future health problems, health-care utilization, and mortality (Idler and Benyamini, 1997; Mossey and Shapiro, 1982).

Despite the widespread use of SRH to measure health, and the growing literature examining its measurement properties, we know surprisingly little about how individuals form their health judgment. SRH is a *social* construct (Kaplan et al., 1976) generated by a complex subjective cognitive-emotional process, grounded in context and culture (Jylhä, 2009). The foundation of SRH is the respondents' physiological state and their knowledge of this state, but a multitude of other inputs and influences—such as health expectations, peer comparisons, personality characteristics—likely

* Corresponding author. Social Science Centre, Room 5330, Western University, London, Ontario N6A 5C2, Canada.
E-mail address: anna.zajacova@uwo.ca (A. Zajacova).

http://dx.doi.org/10.1016/j.socscimed.2017.06.019
0277-9536/

play a role. The basic components of SRH include the presence or absence of illnesses and medical conditions, health behaviors, and disabilities or limitations in physical functioning (Jylhä, 2009; Krause and Jay, 1994; Manderbacka, 1998; Peersman et al., 2012; Undén and Elofsson, 2006). Other health-related correlates of SRH include symptoms such as pain (Malmusi et al., 2012) and psychological well-being (Han et al., 1999). The health judgment is also influenced by socio-economic factors, especially respondents' level of education (Idler et al., 1999; Krause and Jay, 1994).

The evaluative process and the specific frames of reference used to form the SRH judgment differ across individuals (Krause and Jay, 1994). These differences can be problematic if they occur systematically across groups because they could bias our understanding of health inequalities across those groups. It is therefore critical that we better understand how respondents from different population groups form their SRH judgment.

Most prior studies that aimed to understand group differences in the SRH judgment studied its predictive validity, specifically, how well SRH correlates with subsequent mortality. Generally SRH predicts mortality better among respondents of higher socioeconomic status (Dowd and Zajacova, 2007; Fiscella and Franks, 1997), non-Hispanic white adults compared to racial and ethnic minority members (Woo and Zajacova, 2016), in more recent birth cohorts (Schnittker, 2005), and younger respondents (Zajacova and Woo, 2016). With respect to gender, the predictive validity of SRH is most often found to be higher for men than women in studies from around the world (Benjamins et al., 2004; Benyamini and Idler, 1999; Deeg and Kriegsman, 2003; Hu et al., 2016; Idler and Benyamini, 1997; Nishi et al., 2012; Spiers et al., 2003) although some studies found no significant difference in SRH-mortality associations between men and women (Bath, 2003; Singh-Manoux et al., 2007). An exception is a finding among older Israeli respondents of a slightly higher predictive validity for women (Benyamini et al., 2003).

However, anchoring SRH in mortality to explore gender differences in the meaning or structure of SRH is problematic given the long-recognized gender paradox (Nathanson, 1975): gender patterns in mortality are different from patterns in morbidity. Women tend to have worse health but live longer than men (Verbrugge, 1985, 1989). Trying to understand gender patterns in *health* judgment by anchoring it in *mortality* may therefore fail to uncover meaningful gender differences. In fact, one explanation for the gender paradox is that women and men may differ in evaluating and reporting their health (Idler, 2003).

Therefore, we need to more directly study the meaning and structure of the SRH in men and women: how do physical health, mental health, health behaviors, and other factors contribute to the overall health judgment? We present a comprehensive gender comparison of the structure of the health judgment in a large, nationally-representative sample of US adults across the adult life span. We contribute to the literature in several ways.

First, we include many health measures ranging from health conditions, depressive symptoms, pain, functioning, health behaviors, and health care utilization. This range is important because individuals use many different inputs in their health judgments (Krause and Jay, 1994). More specific to the gender comparisons, men and women are sometimes thought to place different weights on particular inputs when making their health judgment. Women may be more inclusive, taking into account mild symptoms and chronic conditions as compared to men who focus on serious and life-threatening conditions (Benyamini et al., 2000). Men were found to use physical functioning as well as health behaviors to a greater degree in their health judgment than women (Peersman et al., 2012). Health behaviors are of particular importance as they may play a central role in gender differences in health (Rieker et al., 2010). While studying the health inputs and their relationships with SRH, we also consider the role of social factors. The social dimension is relevant because differential access to health-related resources due to women's lower social status is considered a key reason for gender health disparities (Read and Gorman, 2010). Moreover, social and physiological inputs may interact in complex ways to produce the health judgment (Rieker and Bird, 2005; Vlassoff, 2007). Empirically, a prior study found a wide range of health and social inputs relevant for SRH among Swedish adults, with no meaningful gender differences in how the inputs predicted men's and women's health judgments (Undén and Elofsson, 2006).

Second, we describe patterns in SRH across the adult life span. Examining a wide age range is important because the gender patterns in SRH may depend on the life stage (Read and Gorman, 2010), as do patterns in many health conditions. For instance, cardiovascular diseases (CVD) are more prevalent among men than women at younger ages but this difference disappears among older adults (Rieker et al., 2010). Several studies also suggested that women's health ratings may be worse than men's during earlier stages of adulthood (Ross and Bird, 1994) but this gender difference may attenuate or even disappear at older ages (Gorman and Read, 2006; Verbrugge, 1985). There are also indications that the health conditions most important for forming the health judgment change across age, regardless of gender (Krause and Jay, 1994; Read and Gorman, 2011). We therefore include adults aged 25–84 and examine gender differences separately among young, middle-aged, and older adults.

Finally, we address the possibility that the health judgment may not operate as a smooth linear continuum for men and women: factors that predict excellent or very good SRH may not be the same factors that lead individuals to report poor health (Jylhä, 2009; Kaplan and Barol-Epel, 2003). Yet typical frequentist analyses model only the central tendencies and thus may obscure potential gender differences at the tails of the SRH distribution. We thus conclude our analysis with a proof-of-concept Bayesian approach to exploring gender differences with respect to one health measure −mental health– at specific SRH levels.

2. Data and methods

2.1. Data

We used data from the 2002–2015 National Health Interview Survey (NHIS). NHIS is an ongoing annual cross-sectional survey representative of the civilian non-institutionalized population of the United States (National Center for Health Statistics, 2016). In-person interviews are conducted by U.S. Census interviewers in about 35,000 households every year. The purpose of the NHIS is to monitor the levels, trends, and correlates of health and disability in the American population. We accessed the data via the harmonized Integrated Health Interview Survey version (Minnesota Population Center and State Health Access Data Assistance Center, 2016).

Sample. In the NHIS sampling frame, a random subsample of one adult per household ("sample adult") was asked detailed health questions. Our analytic sample consists of US-born respondents age 25–84 who were asked these detailed health questions in 2002–2015 interviews. We included only US-born adults to minimize potential confounding due to language or cultural issues (Bzostek et al., 2007). The NHIS interview range 2002–2015 was selected because 2002 is the first year when question about arthritis, a potentially-salient highly prevalent condition, was asked systematically; 2015 was the most recent wave available when we conducted the analysis. Among 296,679 US-born “sample adults” age 25–84 interviewed in 2002–2015, 0.07% did not have valid SRH information and were excluded from the analyses. Additional information on missingness is below in the Analytic Strategy section.

2.2. Measures

Self-rated health was assessed using the following question: “would you say your health in general is excellent (1), very good (2), good (3), fair (4), or poor (5)?” It is used as a 5-point variable, specified as continuous in the partial linear models and categorical otherwise. Higher values indicate worse health as is typical in the literature and to be consistent with the health measures where higher values also indicate worse health.

Health measures comprise a range of conditions, functioning indices, and other measures. *Chronic health conditions* are dichotomous responses regarding whether a doctor has ever diagnosed angina pectoris, arthritis, cancer, coronary heart disease, chronic bronchitis, diabetes, emphysema, heart attack, other heart condition, hypertension, kidney disease, liver disease, and stroke. *Acute conditions* are measured with a dichotomous indicator capturing whether the respondent experienced a head or chest cold or ‘stomach problems’ in the past two weeks. *Pain* is a summation index ranging from 0 to 4 of positive answers (=1) regarding whether the respondent experienced, in the previous three months, lower-back pain, neck pain, facial pain, or severe headaches or migraines. *The Kessler psychological distress (K6) scale* (Kessler et al., 2002) is an index (range 0–24) summed from responses to six questions regarding symptoms experienced over the past 30 days. Respondents were asked how often they felt nervous, hopeless, worthless, restless/fidgety, so sad that nothing could cheer him/her up, and that everything was an effort. For each symptom, respondents answered all of the time (4), most of the time (3), some of the time (2), a little of the time (1), or none of the time (0). *Functional limitations* is an index (range 0–36) comprising nine questions about difficulties performing the following activities without special equipment: carry 10 pounds, walk up 10 steps, grasp objects, reach overhead, sit two hours, stand two hours, stoop/bend/kneel, walk ¼ mile, and push large objects. Each activity was coded as not at all difficult (0), only a little difficult (1), somewhat difficult (2), very difficult (3), and cannot do or does not do (4). *Any activity limitation* is a measure of disability; it is an IHIS-created dichotomous variable coded 1 if a respondent needs help with, or is limited in, any of a range of activities of daily life including ADLs, IADLs, and work limitations. The variable “*10 + doctor visits*” identifies respondents who received health care 10 or more times during the past 12 months. *Hospital stay* captures persons with an overnight hospital stay during the past 12 months, excluding childbirth. *Any bed days* is a dichotomous indicator whether, during the preceding 12 months, an illness or injury kept the respondent in bed for more than half a day. Two variables capture health behaviors: *body mass index (BMI)*, a continuous variable, and a dichotomous indicator whether the respondent is a *current smoker*.

Covariates included in all models are age (continuous) and gender (male = reference). Additional variables are auxiliary and used only in Table 2 to capture respondents' socioeconomic status: education (less than high school, high school diploma, some college, bachelor's degree or more), marital status (married or cohabiting, divorced or separated, widowed, and never married), employment status (employed, unemployed, not in labor force, retired, other), household income (top, middle, and bottom income tertile), and owning versus renting one's home. The detailed coding and distribution of these characteristics in the sample is not shown due to space constraints but is available from the authors.

Table 1
Select characteristics of the analytic sample, NHIS 2002–2015, by gender.

	Male	Female	Gender Δ
Self-rated health			***
Excellent	26.95%	25.23%	
Very good	32.90%	32.72%	
Good	26.37%	27.11%	
Fair	10.22%	11.22%	
Poor	3.56%	3.72%	
Health conditions			
Acute condition	14.24%	15.96%	***
Arthritis	23.32%	31.84%	***
Angina	3.10%	2.40%	***
Cancer	8.98%	11.00%	***
CHD	7.05%	3.91%	***
Chronic bronchitis	3.32%	6.50%	***
Diabetes	11.36%	10.75%	***
Emphysema	2.63%	2.17%	***
Heart attack	5.51%	2.91%	***
Other heart condition	8.63%	9.25%	***
Hypertension	34.23%	33.82%	
Kidney disease	1.87%	1.99%	*
Liver condition	1.69%	1.58%	*
Stroke	3.11%	3.24%	
Physical functioning			
Activity imitation	18.04%	20.14%	***
Functional limitation	33.76%	43.75%	***
Limitation index	2.86 (0.03)	4.49 (0.03)	***
Other health measures			
Received care 10 + times	12.03%	17.42%	***
Overnight hospital	8.67%	11.79%	***
Any bed days	33.14%	41.51%	***
Pain index (0–4 range)	0.56 (0.00)	0.78 (0.00)	***
CES-D score (0–24 range)	2.22 (0.02)	2.81 (0.02)	***
Health behaviors			
Current smoker	23.82%	19.82%	***
BMI	28.26 (0.02)	27.70 (0.03)	***
Age	49.81 (0.08)	51.28 (0.08)	***
Proportion of the sample	45.2%	54.8%	–

*p < 0.05, **p < 0.01, ***p < 0.001.
Adjusted for NHIS complex sample design. For continuous variables, mean (standard error) is shown.
Descriptives for the total sample and for socioeconomic covariates used in Table 2 are available on request.

Table 2
Gender difference in SRH, net of different covariate groups.

	Age 25-44	Age 45-64	Age 65-84
Model set 1: Demographics	1.12***	0.99	0.92***
Model set 2: Demographics, SES	1.04**	0.90***	0.84***
Model set 3: Demographics, health	1.01	0.84***	0.75***
Model set 4: Demographics, health, SES	1.00	0.83***	0.75***

*p < 0.05, **p < 0.01, ***p < 0.001.
Adjusted for NHIS complex sampling design.
Each cell shows odds ratio for “female” from a separate model of SRH with different covariate groups. For SRH, higher values = worse health, so OR>1 indicates that women's SRH is worse than men's.
Demographic covariates include age, race/ethnicity, and region of residence.
Health measures in include the full set of health variables as shown in Table 1.
SES measures include education, marital status, employment status, household income, and owning versus renting one's home. Details on coding and distribution of these variables available on request.

2.3. Analytic strategy

First, descriptive analyses summarize gender differences in SRH and all health measures. Next we estimate a series of ordered logistic models of SRH on gender, net of select covariate groups, in order to assess gender differences in SRH levels. The models are estimated separately for three age groups: 25–44, 45–64, 65–84. We also visualize the average SRH for men and women across the life span with and without health and social covariates. The line graphs are based on estimates from partial linear models (Lokshin, 2006) of SRH. In the models, age is a nonparametric component estimated via Stata's lowess procedure and all other predictors are entered as linear (parametric) components.

The main research question is answered using ordered logistic models of SRH where gender and health measures are interacted. The interaction effects test whether men and women differ in how health measures predict their SRH.

The above analyses model the central tendencies (mean SRH in the partial linear models) or assume proportional odds (ordered logistic models), effectively forcing the effect of all predictors —including the gender comparisons– to be equal across all levels of SRH. What if this is not the case? For instance, perhaps respondents of one gender are better at distinguishing excellent from very good health while the others are better at distinguishing fair from poor (c.f., Jylhä, 2009; Kaplan and Barol-Epel, 2003). To examine this, we add a simple but novel proof-of-concept Bayesian analysis to study the association of a health measure with SRH at each SRH level separately. For the health measure, we chose the K6 score for substantive and methodological reasons: substantively, mental health is a fundamental dimension that can influence health evaluation more broadly than other measures (Han and Jylha, 2006); methodologically this variable lent itself well to approximating using the Poisson distribution. We also analyzed the sum of chronic conditions and physical limitations index and found comparable results (available on request). We estimate a regression model of K6 as a function of SRH level as a set of dummy variables, controls, and their interactions with gender. Note that unlike in the frequentist analyses above where SRH is the outcome, here the health measure is the dependent variable and is estimated using an appropriate GLM: Poisson model was appropriate for the K6 distribution. From the regression model we generate posterior predictive distributions of K6 values at each SRH level and gender at mean levels of all controls, obtained by integrating over the posterior distributions of the regression coefficients. We then calculate the proportions of distributions of K6 at adjacent SRH levels that are 'correctly aligned' – that is, the proportion of K6 scores for a given SRH level that are worse or equal to K6 scores for a better adjacent SRH level. For instance, when comparing with fair and good SRH, correct alignment captures the proportion of K6 values at the fair SRH that are worse or equal to K6 values at the good SRH level. These 'correct alignment' proportions can be then compared for men and women to explore whether the genders differ in how a health measure and SRH are associated at different levels of the health judgment. We use vague (noninformative) priors for the Poisson coefficients (Gelman et al., 2013), Normal (1,1000). We estimate the models using JAGS with 3 MCMC chains, 10,000 iterations, and burn-in of 5000; we thin the chains taking every third sample.

We also conducted extensive auxiliary analyses to explore the robustness of findings. (1) We calculated correlation matrices of SRH and all health measures. This step confirmed a lack of collinearity problems among the individual health measures in the regression models: on average, the correlations among the health indices were $r = 0.18$, with only three out of 55 correlation coefficients exceeding $r = 0.4$. (2) Because the SRH evaluation process could have changed in the population over the course of the 14 interview years (Schnittker and Bacak, 2014), we also estimated models stratified by interview year, separately for 2002–2008 and 2008–2015. (3) The comparison of coefficients from nested or non-nested nonlinear models such as the ordered logit we use is problematic because the effects are a function not only of the included versus omitted variables, but additionally by the rescaling of the model, meaning they are influenced by the error variance, which in turn is a function of the explained variance (Aaron, 1967; Kohler et al., 2011; Mood, 2010). We took care never to directly compare the coefficients across different models in our interpretation. Additionally, we re-estimated all models using linear regression, that is, specifying the SRH as a continuous outcome. 4) To evaluate the stability of the main finding regarding the gender similarity in the effect of health measures on SRH, we estimated the model with and without socioeconomic controls, included additively or interacted with gender. 5) The missingness on key variables in our analytic sample was low. It was 0% for age and gender, less than 0.3% for nearly all health conditions, 0.6% for education. The highest missingness was 1.7% for K6 and 4% for BMI. We re-estimated all models in two ways. We used listwise deletion in one set of results and we also estimated multiply-imputed models.

The findings from all auxiliary analyses were equivalent to the results shown below. All tables are available from the authors.

3. Results

Table 1 shows the characteristics of the analytic sample. The mean age is about 50 years and approximately 55% are women. As expected, women are somewhat less likely to report excellent or very good health and more likely to report fair or poor health. Men report more cardiovascular problems while women report more conditions like arthritis or acute illness, more pain, higher depressive symptoms, and more health care utilization. For hypertension and stroke there is no gender difference.

Table 2 summarizes gender differences in SRH levels. Model set 1 controls for demographic characteristics: age, race/ethnicity, and region of residence. The $OR = 1.12$ ($p < 0.001$) indicates that women age 25–44 have 12% higher odds of reporting a worse SRH category compared to men, net of the demographics. At age 45–64, there is no gender difference ($OR = 0.99$) and among older adults, women report significantly better SRH than men ($OR = 0.92$, $p < 0.001$). Taking into account socioeconomic factors (education, marital status, employment status, household income, and home ownership), the odds ratios are 1.04 ($p < 0.01$), 0.90 ($p < 0.001$), and 0.84 ($p < 0.001$) at younger through older ages, respectively. Net of health measures in Model set 3, women's SRH is comparable to men's at ages 25–44 and significantly better at 45–64 ($OR = 0.84$, $p < 0.001$) and 65–84 ($OR = 0.75$, $p < 0.001$). The further inclusion of socioeconomic factors in Model set 4 yields similar odds ratios.

Fig. 1a shows mean SRH by age for men and women across the adult life span. Unsurprisingly, the average SRH becomes progressively worse (higher) with age for both genders. Corroborating the results from Model set 1 in Table 2, women's SRH is worse than men's at younger ages but there is a crossover in mid adulthood such that among adults in their 60s and older, women's SRH is slightly better than men's. Fig. 1b then visualizes the SRH-age pattern net of health measures. There are two things to note. First, the gender pattern is similar to the unadjusted findings in that women's SRH is slightly worse through mid-adulthood but then becomes slightly better (significance not tested but the size of the difference appears small). Second, the age pattern in SRH is much flatter, especially in older ages. This indicates that the increases in SRH with age can be largely explained by increasing health problems.

Table 3 shows results for the main research question. There are three important findings to note. First, with just a few exceptions, all health measures are significantly associated with SRH in all three age groups (among men), in the expected direction and often with large magnitude of effects (i.e., diabetes increases the odds of a worse SRH rating 1.8 times in older men and 2.8 times in younger men). Second, with some exceptions, women and men are similar

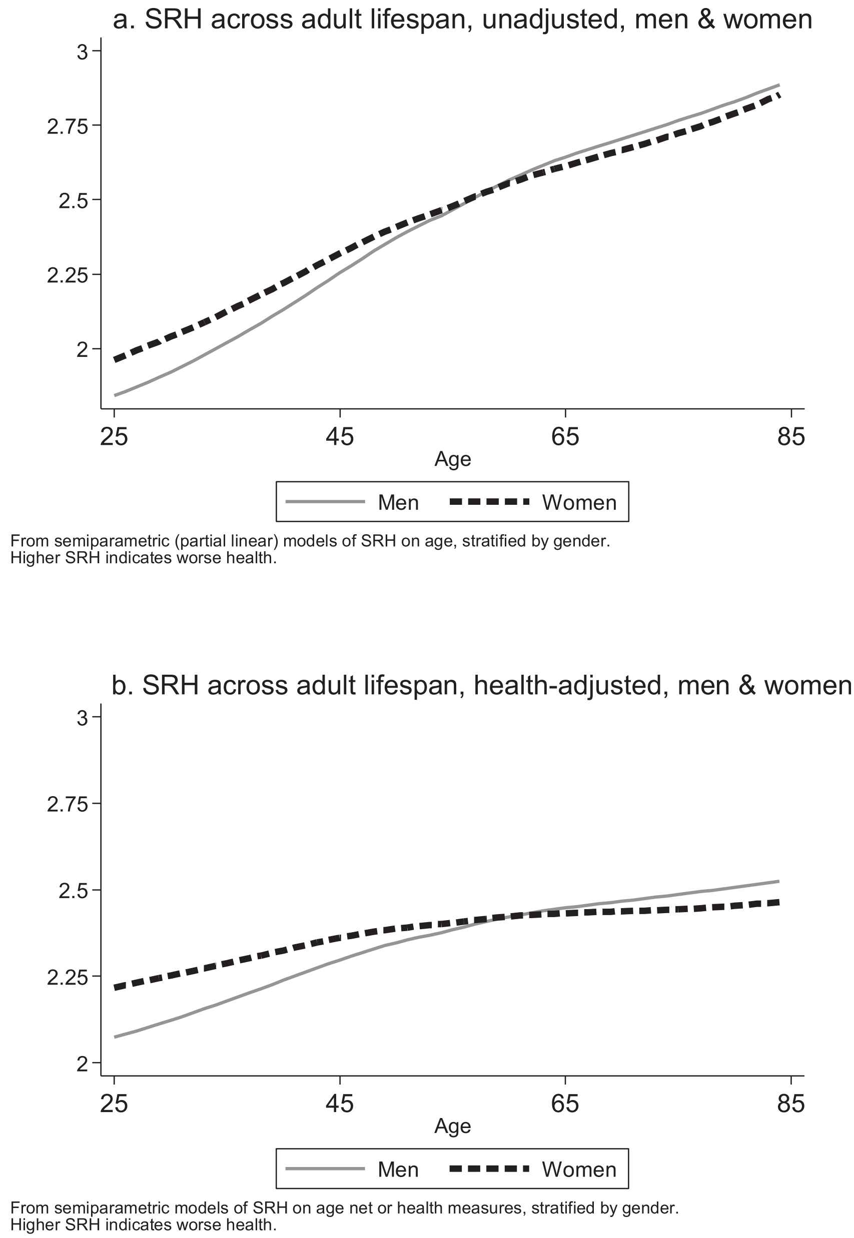

Fig. 1. (a)SRH across adult lifespan, unadjusted, men & women. (b) SRH across adult lifespan, health-adjusted, men & women.

in how health measures and their SRH are linked: the interaction effects are not significant. The notable exceptions are physical functioning and smoking, as well as smoking-related conditions like chronic bronchitis and emphysema, which predict SRH significantly more strongly among men than women. For example, the effect of current smoking on SRH is 16% weaker for women than men at age 45–64 and 23% weaker among women at ages 65–84.

Can we infer that the findings above, focused on gender comparison of the mean SRH, apply at all levels of SRH? Table 4 shows how well adjacent pairs of SRH levels distinguish the distributions of one health measure, K6. We show the probability that in randomly chosen pairs of simulated respondents at adjacent SRH levels, the respondent with higher SRH (worse health) also has a higher or equal K6 score (worse or equal psychological distress symptoms). This probability represents a 'correct alignment'. If SRH predicted the distributions of K6 perfectly, then 100% of random simulated observations pairs would be 'correctly aligned.' The first number in the table, 0.47 for men, indicates that if we compare men in very good versus excellent health, 47% of men in very good health will have an equal or worse K6 score compared to men in excellent health. Two patterns are noteworthy. At worse SRH levels, the K6 distributions become better separated, indicating a higher validity of the health judgment: at poor versus fair SRH, 75% and 79% of pairs of simulated observations are aligned correctly among men and women, respectively. The difference between men and women is modest relative to the pattern across SRH levels. The gender similarity can also be observed in Fig. 2 which shows the simulated K6 distribution overlaps for two pairs of SRH levels: excellent versus very good and fair versus poor. The patterns at both ends of the scale are similar for men and women.

4. Discussion

This study explored how US men and women across the adult

Table 3
Ordered logit models of SRH, with health by gender interactions.

	Age 25-44	Age 45-64	Age 65-84
Female	1.11	1.03	1.00
Age	1.01***	1.00	1.01***
Main effects of health anchors			
Angina	1.71**	1.15*	1.11
Arthritis	1.30***	1.11***	0.96
Cancer	1.47***	1.23***	1.12***
CHD	1.91***	1.61***	1.47***
Chronic bronchitis	1.25***	1.28***	1.44***
Diabetes	2.85***	2.26***	1.83***
Emphysema	1.92**	2.40***	2.22***
Heart attack	1.43*	1.38***	1.33***
Other heart condition	1.30***	1.17***	1.25***
Hypertension	1.61***	1.68***	1.41***
Kidney disease	1.95***	1.62***	1.74***
Liver disease	1.79***	1.70***	1.50***
Stroke	1.58**	1.49***	1.29***
Pain	1.19***	1.15***	1.13***
Acute conditions in past 2 weeks	1.07*	1.08**	1.10*
K6 depression score	1.08***	1.07***	1.08***
Functional limitations	1.08***	1.09***	1.10***
Any activity limitation	3.06***	2.89***	2.13***
BMI	1.07***	1.04***	1.01***
Current smoker	1.89***	2.11***	1.58***
10 + doctor visits	1.39***	1.36***	1.39***
Hospital stay	1.30***	1.24***	1.17***
Any bed days	1.15***	1.10***	1.10**
Interaction terms with female			
Angina	0.73	0.86	1.02
Arthritis	1.01	1.01	1.06
Cancer	1.06	0.92	0.96
CHD	1.29	0.99	1.04
Chronic bronchitis	0.99	0.95	0.81*
Diabetes	0.87	0.96	1.02
Emphysema	0.85	0.89	0.81*
Heart attack	0.94	0.92	0.95
Other heart condition	0.92	0.93	1.02
Hypertension	1.00	0.99	1.07
Kidney disease	1.04	1.01	0.97
Liver disease	0.99	1.05	1.08
Stroke	1.02	0.93	1.00
Pain	0.97	0.98	1.00
Acute conditions in past 2 weeks	1.09*	1.04	1.07
K6 depression score	1.00	1.00	1.00
Functional limitations	1.00	0.99*	0.99*
Any activity limitation	0.98	0.87**	0.75***
BMI	1.00	1.00	0.99
Current smoker	0.98	0.84***	0.77***
10 + doctor visits	0.93	1.02	1.03
Hospital stay	0.82**	0.98	0.92
Any bed days	0.96	1.02	1.02

Adjusted for NHIS complex sampling design.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Table 4
Comparing the values of K6 depression index across adjacent pairs of SRH levels: proportion of observations where a higher SRH level was linked to *greater* (worse) *or equal* K6 values compared to the lower SRH level.

	Men	Women
Excellent - very good	0.47	0.52
Very good – good	0.56	0.60
Good – Fair	0.76	0.76
Fair - poor	0.75	0.79

life span evaluate their health: Do they differ in which health measures drive their SRH judgment? This is an important issue because we frequently assess gender differences in health using the SRH measure, and if men and women differ in how they generate their SRH assessment, then our understanding of the gender health differences could be biased. Several theoretical developments suggested that men and women may indeed differ when evaluating their health (Schneider et al., 2012), and empirical research documented a higher predictive power of SRH for mortality for men than women (Benyamini and Idler, 1999; Idler and Benyamini, 1997).

Our central finding, however, is that men and women are remarkably similar in how they incorporate a wide range of chronic and acute health conditions, functioning, health-care utilization, and health behaviors in their SRH evaluation. We used a large, nationally-representative sample across the adult life span with plentiful statistical power to detect gender differences in the associations between a range of health measures and SRH but found only isolated differences. This gender similarity in the SRH structure among US men and women corroborates findings among adults from Sweden (Undén and Elofsson, 2006) where the authors found the association of health and social correlates with SRH “surprisingly similar” for men and women. Our findings also echo Case and Paxson's (2005) study using older NHIS data where chronic conditions predicted SRH similarly among men and women, as well as findings from Spain where men's and women's SRH ratings were similar once differences in the prevalence of health conditions were taken into account (Malmusi et al., 2012).

The gender differences we found in our sample were concentrated among older respondents, where two domains played a larger role in men's health judgments than in women's judgements: activity limitations, and smoking-related behaviors (being a current smoker) and conditions, to a lesser degree (emphysema and chronic bronchitis). These two domains of gender differences are reminiscent of Peersman et al. (2012) study where Belgian men were more likely than women to use physical functioning and health behaviors as a frame of reference. The similarity of findings across US and European studies suggests that our results are generalizable beyond the US or English-speaking population.

Our finding that men and women are similar in *concurrent* validity of SRH in terms of comparable health-related frames of reference are contrary to the *predictive* validity of SRH, which is typically found to be higher for men. While it's not obvious why the gender patterns are different in mortality versus morbidity, it fits the well-known morbidity/mortality gender paradox wherein gender patterns observed for mortality differ from those observed for morbidity (Nathanson, 1975). Perhaps there are differences in how men and women incorporate particularly serious conditions or in the severity of the conditions (Case and Paxson, 2005; Verbrugge, 1989). Some evidence for this possibility is that men weighted smoking-related problems and activity limitations more heavily than women, especially at older ages when these factors are likely to predict rapidly declining health.

Additional discoveries are informative for our understanding of the health evaluation process. Somewhat surprisingly, nearly all of the 24 health conditions, problems, and behaviors we included in our study were significant independent predictors of the SRH judgments in all age groups. This indicates that respondents consider numerous dimensions of health when they form their SRH judgment. Functioning is paramount: being limited in work or activities of daily living more than doubled the odds of reporting worse SRH among men at all ages and having emphysema, which makes physical activity difficult, doubled the odds or worse health among middle-aged and older adults. Interestingly, a diagnosis of diabetes is particularly impactful among young and middle-aged adults, perhaps because the respondents worry about future health problems due to diabetes. In contrast, arthritis is not a significant predictor of SRH among older men, perhaps because its ubiquity at older ages makes it a less salient condition when other health measures are considered (on its own, arthritis predicts

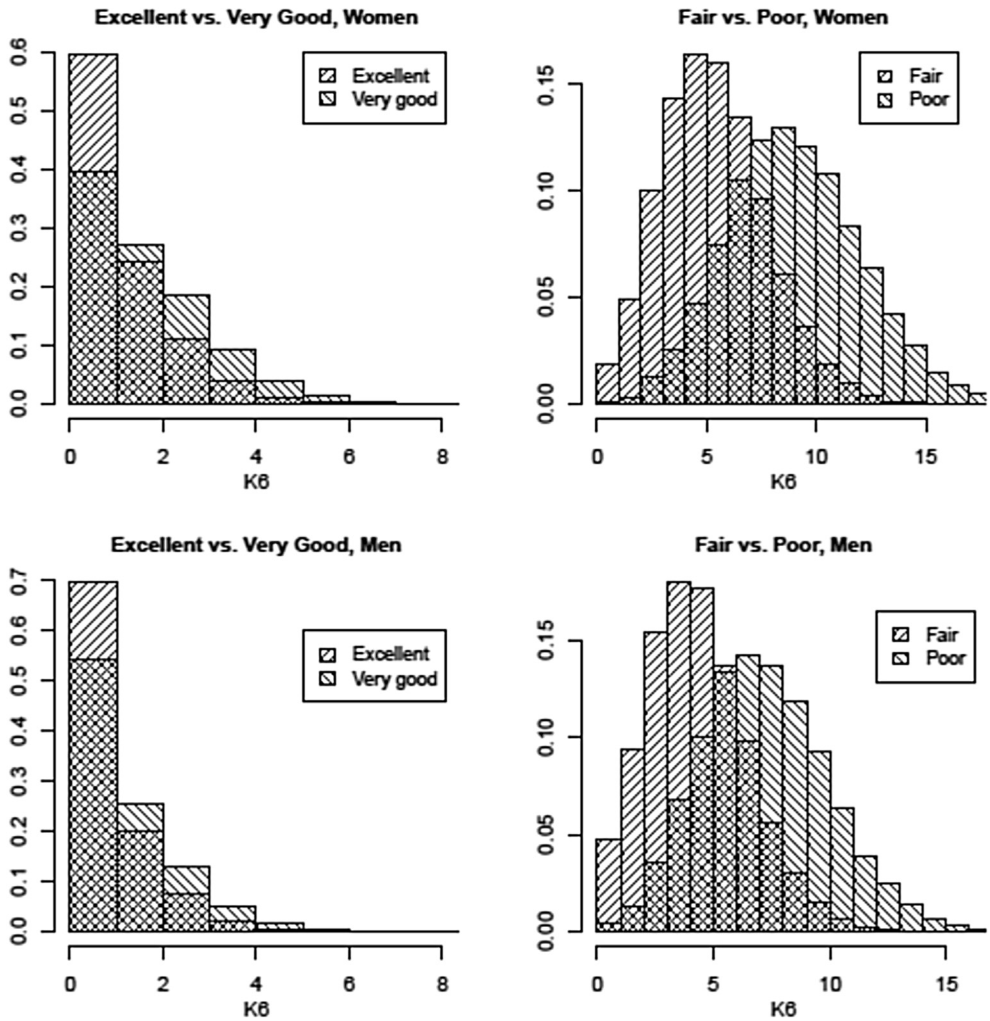

Fig. 2. Distribution of K6 scores at adjacent levels of SRH for men and women. Data for the plots come from posterior predictive distributions of K6 scores, estimated using a Poisson regression of K6 scores as a function of SRH, gender, age, and their interactions.

worse SRH in all age groups).

As expected, the average health rating become worse with age — but a large part of the age trend is explained by the increasing number of health problems after mid-adulthood. In other words, net of health problems, the age pattern in SRH is fairly flat at older ages (cf. Idler, 1993). This pattern further corroborates the high validity of SRH: the age-related changes in SRH are largely due to the increasing prevalence of the underlying health conditions.

One additional result regarding the SRH levels is noteworthy. In agreement with prior literature, we found that women's health ratings are worse on average than men's (Rieker and Bird, 2005). The gender comparison, however, differs across age: the female disadvantage is largest in young adulthood and diminishes gradually with age to a point of a crossover in mid-adulthood, so that among older adults, women report better SRH than men (McCullough and Laurenceau, 2004; Read and Gorman, 2011). This gross age pattern of gender differences is sensitive to the inclusion of covariates in one important respect: taking health measures into account, in particular, changes the relative SRH levels so that at younger ages, women have SRH levels comparable to men, while in middle and older adulthood, women have better SRH. A more general summary of the gender comparison in SRH across age, therefore, is that men experience a steeper worsening in SRH across age than women do.

These complex age-dependent patterns are clearly influenced by multiple factors. At older ages, men may experience more, or more severe, health problems than women, perhaps in part due to men's poorer health behaviors earlier in the lifecourse. For instance, smoking and smoking-related conditions figured more prominently in older men's SRH judgment than in women's. Additionally, health ratings are a social construct (Jylhä, 2009) and women's poorer health at younger ages may reflect their more disadvantaged social position (Arber and Cooper, 1999; Vlassoff, 2007). As they age, perhaps women's social roles may become less demanding as their children reach adulthood, potentially freeing women's time that they can invest in their health or lessening stress due to demands of multiple social roles, which could improve women's health perceptions. While the reasons for the steeper decline in men's health remain unknown, the pattern reinforces recent exhortations that we consider "the importance of age in defining men's and women's health status" to avoid overgeneralizations (Read and Gorman, 2011).

Finally, we extended the analysis beyond the focus on central tendencies. Standard frequentist models either focus on mean SRH or depend on the assumption of proportional odds (for ordered logistic models of SRH), effectively forcing the effect of any predictor to be the same at poor health as it is at excellent health. We proposed a basic Bayesian approach that allows us to compare men and women in how well a given health measure (we chose the depressive symptom score because of its centrality to health perceptions) differs between pairs of adjacent SRH categories. This extension corroborated the findings above: men and women were roughly similar in how K6 and SRH were linked, whether at low SRH levels or high; women's SRH levels discriminated between the K6 scores a bit better, perhaps. This approach could be useful in other studies of criterion validity of general health measures (Kaplan et al., 1976).

We note several limitations of this study. First, all health measures are self-reported, which means that any potential sex differences that influence the SRH judgment may, in some form, influence the reporting of the health measures as well. This limitation can be dealt with in future studies by using objectively measured outcomes, such as biological risk markers or physical performance, or health conditions collected from patient records. Second, the health measures we used are only a subset of those that likely matter for the health judgment; moreover, we miss information about important dimensions such as severity or duration. Third, fourteen of the 24 health measures are doctor-diagnosed conditions. To the degree that men and women differ in medical-care seeking behavior (Bertakis et al., 2000) and may be diagnosed differentially (i.e., Maas and Appelman, 2010), these measures could be biased. However, the patterns are also comparable for the ten measures that do not depend on a doctor's diagnosis, which assuages worries about this type of bias. Finally, we interpreted the time-dependent patterns as age – however, there is a strong correlation between age and birth cohort. It is quite possible there are cohort changes in how the SRH judgment is formed, and future studies should focus on determining which of these two time dimensions plays the key role because the implications for future patterns in SRH by gender will differ accordingly.

These limitations notwithstanding, our study showed that men and women incorporate a wide range of health measures into their health judgment – in fact, all 24 health measures we studied were significant predictors of the health judgment. Moreover, men and women across the full adult life span weigh these measures quite similarly in their health judgment. This suggests a common understanding of the importance of different dimensions of health for overall heath rating in both genders and thus provides further corroboration of the high validity of the SRH measure in population-health research.

Acknowledgements

This work was supported in part by the Eunice Kennedy Shriver National Institute of Child Health & Human Development (grants P2CHD047879 and P2CHD058486) and the National Institute of General Medical Sciences (1U54GM104944-01A1) of the National Institutes of Health.

References

Aaron, Antonovsky, 1967. Social class, life expectancy and overall mortality. Milbank Meml. Fund Q. 45 (2), 31–73. http://dx.doi.org/10.2307/3348839.

Arber, Sara, Cooper, Helen, 1999. Gender differences in health in later life: the new paradox? Soc. Sci. Med. 48 (1), 61–76. http://dx.doi.org/10.1016/S0277-9536(98)00289-5.

Bath, Peter A., 2003. Differences between older men and women in the self-rated health–mortality relationship. Gerontologist 43 (3), 387–395. http://dx.doi.org/10.1093/geront/43.3.387.

Benjamins, Reindl, Maureen, Hummer, Robert A., Eberstein, Isaac W., Nam, Charles B., 2004. Self-reported health and adult mortality risk: an analysis of cause-specific mortality. Soc. Sci. Med. 59 (6), 1297–1306.

Benyamini, Yael, Idler, Ellen L., 1999. Community studies reporting association between self-rated health and mortality. Res. Aging 21, 392–401.

Benyamini, Yael, Leventhal, Elaine A., Leventhal, Howard, 2000. Gender differences in processing information for making self-assessments of health. Psychosom. Med. 62 (3), 354–364.

Benyamini, Yael, Blumstein, Tzvia, Lusky, Ayala, Modan, Baruch, 2003. Gender differences in the self-rated health–mortality association: is it poor self-rated health that predicts mortality or excellent self-rated health that predicts survival? Gerontologist 43 (3), 396–405. http://dx.doi.org/10.1093/geront/43.3.396.

Bertakis, Klea D., Azari, Rahman, Jay Helms, L., Callahan, Edward J., Robbins, John A., 2000. Gender differences in the utilization of health care services. J. Fam. Pract. 49 (2), 147, 47.

Bzostek, Sharon, Goldman, Noreen, Pebley, Anne, 2007. Why do hispanics in the USA report poor health? Soc. Sci. Med. 65 (5), 990–1003. http://dx.doi.org/10.1016/j.socscimed.2007.04.028.

Case, Anne, Paxson, Christina, 2005. Sex differences in morbidity and mortality. Demography 42 (2), 189–214.

Cousins, Sandra O'Brien, 1997. Validity and reliability of self-reported health of persons aged 70 and older. Health Care Women Int. 18 (2), 165–174.

Deeg, Dorly J.H., Kriegsman, Didi M.W., 2003. Concepts of self-rated health: specifying the gender difference in mortality risk. Gerontologist 43 (3), 376–386. http://dx.doi.org/10.1093/geront/43.3.376.

DeSalvo, K.B., Bloser, N., Reynolds, K., 2006. Mortality prediction with a single general self-rated health question. A meta-analysis. J. Gen. Intern Med. 21 http://dx.doi.org/10.1111/j.1525-1497.2005.00291.x.

Dowd, Beam Jennifer, Zajacova, Anna, 2007. Does the predictive power of self-rated health for subsequent mortality risk vary by socioeconomic status in the us? Int. J. Epidemiol. 36 (6), 1214–1221. http://dx.doi.org/10.1093/ije/dym214.

Fiscella, Kevin, Franks, Peter, 1997. Does psychological distress contribute to racial and socioeconomic disparities in mortality? Soc. Sci. Med. 45 (12), 1805–1809.

Gelman, Andrew, Carlin, John B., Stern, Hal S., Dunson, David B., Vehtari, Aki, Rubin, Donald, 2013. Bayesian Data Analysis, Third Edition. Chapman and Hall/CRC, Boca Raton, FL.

Gold, Marthe, Franks, Peter, Erickson, Pennifer, 1996. Assessing the health of the nation: the predictive validity of a preference-based measure and self-rated health. Med. Care 34 (2), 163–177.

Gorman, Bridget K., Read, Jen'nan Ghazal, 2006. Gender disparities in adult health: an examination of three measures of morbidity. J. Health Soc. Behav. 47 (2), 95–110. http://dx.doi.org/10.2307/30040304.

Han, B., Jylha, M., 2006. Improvement in depressive symptoms and changes in self-rated health among community-dwelling disabled older adults. Aging & Ment. Health 10 (6), 599–605.

Han, Beth, Small, Brent J., Haley, William E., 1999. The effects of race, gender, and education on the structure of self-rated health among community-dwelling older adults. Ann. N. Y. Acad. Sci. 896 (1), 442–447.

Hu, Yu-Ning, Chen, Pei-Chun, Hsu, Chien-Chi, Yu, Hui-Kung, Chien, Kuo-Liong, Li, Chin-Ching, Hu, Gwo-Chi, 2016. Age and gender differences in the relationship between self-rated health and mortality among middle-aged and elderly people in taiwan—results of a national cohort study. Int. J. Gerontology. http://dx.doi.org/10.1016/j.ijge.2014.08.005. Available online May 31 2016.

Idler, Ellen L., 1993. Age differences in self-assessments of health: age changes, cohort differences, or survivorship? J. Gerontology 48 (6), S289–S300.

Idler, Ellen L., Benyamini, Yael, 1997. Self-rated health and mortality: a review of twenty-seven community studies. J. Health Soc. Behav. 38 (1), 21–37.

Idler, Ellen L., Hudson, Shawna V., Leventhal, Howard, 1999. The meanings of self-ratings of health - a qualitative and quantitative approach. Res. Aging 21 (3), 458–476.

Idler, Ellen L., 2003. Discussion: gender differences in self-rated health, in mortality, and in the relationship between the two. Gerontologist 43 (3), 372–375. http://dx.doi.org/10.1093/geront/43.3.372.

Jylha, Marja, Volpatob, Stefano, Guralnik, Jack M., 2006. Self-rated health showed a graded association with frequently used biomarkers in a large population sample. J. Clin. Epidemiol. 59 (5), 465–471.

Jylhä, Marja, 2009. What is self-rated health and why does it predict Mortality? Towards a unified conceptual model. Soc. Sci. Med. 69 (3), 307–316. http://dx.doi.org/10.1016/j.socscimed.2009.05.013.

Kaplan, Giora, Barol-Epel, Orna, 2003. What lies behind the subjective evaluation of health status? Soc. Sci. Med. 56 (8), 1669–1676.

Kaplan, Robert M., Bush, J.W., Berry, Charles C., 1976. Health status: types of validity and the index of well-being. Health Serv. Res. 11 (4), 478–507.

Kessler, R.C., Andrews, G., Colpe, L.J., et al., 2002. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol. Med. 32 (06), 959–976. http://dx.doi.org/10.1017/S0033291702006074.

Kohler, Ulrich, Bernt Karlson, Kristian, Holm, Anders, 2011. Comparing coefficients of nested nonlinear probability models. Stata J. 11 (3), 420–438.

Krause, Neal M., Jay, Gina M., 1994. What do global self-rated health items measure? Med. Care 32 (9), 930–942.

Lokshin, Michael, 2006. Semi-parametric difference-based estimation of partial linear regression models. Stata J. 6 (3), 377–383.

Lundberg, O., Manderbacka, K., 1996. Assessing reliability of a measure of self-rated health. Scand. J. Soc. Med. 24 (3), 218–224.

Maas, Ahem, Appelman, Y.E.A., 2010. Gender differences in coronary heart disease. Neth. Heart J. 18 (12), 598–602.

Malmusi, Davide, Artazcoz, Lucía, Benach, Joan, Borrell, Carme, 2012. Perception or real Illness? How chronic conditions contribute to gender inequalities in self-rated health. Eur. J. Public Health 22 (6), 781–786. http://dx.doi.org/10.1093/eurpub/ckr184.

Manderbacka, Kristiina, 1998. Examining what self-rated health question is

understood to mean by respondents. Scand. J. Soc. Med. 26 (2), 145–153.
McCullough, Michael E., Laurenceau, Jean-Philippe, 2004. Gender and the natural history of self-rated health: a 59-year longitudinal study. Health Psychol. 23 (6), 651–655.
Minnesota Population Center and State Health Access Data Assistance Center, 2016. Integrated Health Interview Series: Version 6.21. University of Minnesota, Minneapolis. http://www.ihis.us.
Mood, Carina, 2010. Logistic regression: why we cannot do what we think we can do, and what we can do about it. Eur. Sociol. Rev. 26 (1), 67–82. http://dx.doi.org/10.1093/esr/jcp006.
Mossey, Jana M., Shapiro, Evelyn, 1982. Self-rated health: a predictor of mortality among the elderly. Am. J. Public Health 72 (8), 800–808.
Nathanson, Constance A., 1975. Illness and the feminine role: a theoretical review. Soc. Sci. Med. 9 (2), 57–62. http://dx.doi.org/10.1016/0037-7856(75)90094-3.
National Center for Health Statistics, 2016. National Health Interview Survey. Retrieved June 17, 2016, 2016. http://www.cdc.gov/nchs/nhis/.
Nishi, Akihiro, Kawachi, Ichiro, Shirai, Kokoro, Hirai, Hiroshi, Jeong, Seungwon, Kondo, Katsunori, 2012. Sex/gender and socioeconomic differences in the predictive ability of self-rated health for mortality. PLoS One 7 (1), 1–7. http://dx.doi.org/10.1371/journal.pone.0030179.
Peersman, Wim, Cambier, Dirk, De Maeseneer, Jan, Willems, Sara, 2012. Gender, educational and age differences in meanings that underlie global self-rated health. Int. J. Public Health 57 (3), 513–523. http://dx.doi.org/10.1007/s00038-011-0320-2.
Read, Jen'nan G., Gorman, Bridget K., 2010. Gender and health inequality. Annu. Rev. Sociol. 36 (1), 371–386. http://dx.doi.org/10.1146/annurev.soc.012809.102535.
Read, Jen'nan G., Gorman, Bridget K., 2011. Gender and health revisited. In: Pescosolido, B.A., Martin, J.K., McLeod, J.D., Rogers, A. (Eds.), Handbook of the Sociology of Health, Illness, and Healing, Handbooks of Sociology and Social Research. Springer, New York, pp. 411–429.
Rieker, Patricia P., Bird, Chloe E., 2005. Rethinking gender differences in health: why we need to integrate social and biological perspectives. Journals Gerontology Ser. B Psychol. Sci. Soc. Sci. 60 (Special Issue 2), S40–S47. http://dx.doi.org/10.1093/geronb/60.Special_Issue_2.S40.
Rieker, Patricia P., Bird, Chloe E., Lang, Martha E., 2010. Understanding gender and health: old patterns, new trends, and future directions. In: Bird, C.E., Conrad, P., Fremont, A.M., Timmermans, S. (Eds.), Handbook of Medical Sociology. Vanderbilt University Press, Nashville, TN, pp. 52–74.
Ross, Catherine E., Bird, Chloe E., 1994. Sex stratification and health lifestyle: consequences for Men's and Women's perceived health. J. Health Soc. Behav. 35 (2), 161–178.
Schneider, Udo, Pfarr, Christian, Schneider, Brit S., Ulrich, Volker, 2012. I feel good! Gender differences and reporting heterogeneity in self-assessed health. Eur. J. Health Econ. 13 (3), 251–265.
Schnittker, Jason, 2005. When mental health becomes health: age and the shifting meaning of self-evaluations of general health. Milbank Q. 83 (3), 397–423.
Schnittker, Jason, Bacak, Valerio, 2014. The increasing predictive validity of self-rated health. PLoS One 9 (1), 1–11.
Singh-Manoux, Archana, Dugravot, Aline, Shipley, Martin J., Ferrie, Jane E., Martikainen, Pekka, Goldberg, Marcel, Zins, Marie, 2007. The association between self-rated health and mortality in different socioeconomic groups in the gazel cohort study. Int. J. Epidemiol. 36 (6), 1222–1228. http://dx.doi.org/10.1093/ije/dym170.
Spiers, Nicola, Jagger, Carol, Clarke, Michael, Arthur, Antony, 2003. Are gender differences in the relationship between self-rated health and mortality Enduring? Results from three birth cohorts in melton mowbray, United Kingdom. Gerontologist 43 (3), 406–411. http://dx.doi.org/10.1093/geront/43.3.406.
Undén, Anna-Lena, Elofsson, Stig, 2006. Do different factors explain self-rated health in men and women? Gend. Med. 3 (4), 295–308. http://dx.doi.org/10.1016/S1550-8579(06)80218-4.
Verbrugge, Lois M., 1985. Gender and health: an update on hypotheses and evidence. J. Health Soc. Behav. 26 (3), 156–182.
Verbrugge, Lois M., 1989. The twain meet: empirical explanations of sex differences in health and mortality. J. Health Soc. Behav. 30 (3), 282–304.
Vlassoff, Carol, 2007. Gender differences in determinants and consequences of health and illness. J. Health, Popul. Nutr. 25 (1), 47–61.
Woo, Hyeyoung, Zajacova, Anna, 2016. Predictive strength of self-rated health for mortality risk among older adults in the United States: does it differ by race and ethnicity? Res. Aging. http://dx.doi.org/10.1177/0164027516637410.
Zajacova, Anna, Dowd, Jennifer Beam, 2011. Reliability of self-rated health in us adults. Am. J. Epidemiol. 174 (8), 977–983. http://dx.doi.org/10.1093/aje/kwr204.
Zajacova, Anna, Woo, Hyeyoung, 2016. Examination of age variations in the predictive validity of self-rated health. Journals Gerontology Soc. Sci. 71 (3), 551–557.