Social Science & Medicine 228 (2019) 41–50

Contents lists available at ScienceDirect

Social Science & Medicine

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

Is the story about sensitive women and stoical men true? Gender differences in health after adjustment for reporting behavior

Anna Oksuzyan[a,*], Maciej J. Dańko[a,1], Jennifer Caputo[a], Domantas Jasilionis[a,b], Vladimir M. Shkolnikov[a,c]

[a] Max Planck Institute for Demographic Research, Konrad-Zuse Str.1, 18057, Rostock, Germany
[b] Demographic Research Centre, Vytautas Magnus University, Jonavos Str. 66, 44138, Kaunas, Lithuania
[c] National Research University Higher School of Economics, Bolshoy Tryokhsvyatitelsky Pereulok 3, 109028, Moscow, Russian Federation

ARTICLE INFO

Keywords:
Gender
Self-rated health
Reporting behavior
Reporting heterogeneity
Gender stereotypes

ABSTRACT

Research indicates that women have higher levels of physical disability and depression and lower scores on physical performance tests compared to men, while the evidence for gender differences in self-rated health is equivocal. Scholars note that these patterns may be related to women over-reporting and men under-reporting health problems, but gender differences in reporting behaviors have not been rigorously tested. Using Wave 1 of the Survey of Health, Ageing and Retirement in Europe (SHARE), the present study investigates the extent to which adjusting for differences in reporting behavior modifies gender differences in general health. We also examine whether men and women's reporting behaviors are consistent across different levels of education. After adjusting for reporting heterogeneity, gender differences in both poor and good health widened. However, we found no clear gender-specific patterns in reporting either poor or good health. Our findings also do not provide convincing evidence that education is an important determinant of general health reporting, although the female disadvantage in poor health and the male advantage in good health were more apparent in lower than higher education groups at all ages. The results challenge prevailing stereotypes that women over-report and men under-report health problems and highlight the importance of attending to health problems reported by women and men with equal care.

1. Background

A substantial body of research has shown that women fare worse on physical tests (Bohannon et al., 2006) and have higher levels of disability, functional limitations (Palacios-Ceña et al., 2012), and depression than their male counterparts (Oksuzyan et al., 2010; Salk et al., 2017). However, gender differences in morbidity are less consistent and appear to vary across chronic conditions: while women suffer more from non-acute disabling conditions, e.g. arthritis, men are more likely to have acute life-threatening conditions, e.g. myocardial infraction (Case and Paxson, 2005; Crimmins et al., 2011; Oksuzyan et al., 2018). Additionally, gender differences in self-rated health (SRH) vary across national contexts. Although women in most European countries (EU) and in the U.S. tend to report poorer health than men (Case and Paxson, 2005; Crimmins et al., 2011; Dahlin and Härkönen, 2013), accounting for differences in socio-demographic characteristics, chronic conditions, and lifestyle behaviors between men and women substantially reduces gender gaps in SRH, and may even reverse women's disadvantage (Crimmins et al., 2011; Dahlin and Härkönen, 2013).

Because women and men provide similar ratings of health at the same levels of morbidity, some scholars have argued that gender inequalities in SRH can be explained by the distribution of non-lethal disabling vs. acute life-threatening conditions among men and women (Case and Paxson, 2005). However, other researchers have suggested that the female disadvantage in SRH results from gender inequalities in social roles, with the expectation that gender differences in health will be smallest in more egalitarian countries and among more educated groups, which are less likely to adhere to traditional gender roles (Bambra et al., 2008; Brewster and Padavic, 2000). Providing support for this hypothesis, UK-based research has shown that gender differences in health are particularly pronounced in socially disadvantaged groups (Cooper, 2002). Within the EU context, scholars have shown that the risk of reporting poor health among women is highest in Southern countries (Italy and Portugal), while no gender differences in

* Corresponding author. Konrad-Zuse-Straße 1, 18057, Rostock, Germany.
E-mail address: oksuzyan@demogr.mpg.de (A. Oksuzyan).
[1] Shared first authorship.

https://doi.org/10.1016/j.socscimed.2019.03.002
Received 19 October 2018; Received in revised form 28 January 2019; Accepted 3 March 2019
Available online 07 March 2019
0277-9536/

SRH are found in Corporatist countries (Belgium, France, and Germany), which have well-compensated and extensive maternity leave (Bambra et al., 2008). However, in Social Democratic countries (Denmark, Sweden, Norway, and the Netherlands), which are widely considered the most progressive countries in terms of gender equality, women have moderately high risk of reporting poor health (Bambra et al., 2008). A possible double burden on women who combine work and family in more egalitarian countries and in higher education groups was suggested to explain these findings.

Another explanation for why women report poorer health than men is that they have greater somatic awareness and are more willing to admit health problems and seek medical advice (Benyamini et al., 2000; Courtenay, 2000). Consistent with this possibility, Macintyre and colleagues (1996) found that the female disadvantage in health was limited to malaise symptoms, such as sleep problems, concentration difficulties, and worrying, while the female excess in morbidity was less apparent or even reversed for physical symptoms. A review study suggested that women were more likely than men to report somatic symptoms whether they were related to medical conditions or medically unexplained (Barsky et al., 2001). Findings of strong associations between SRH and serious diseases in both genders, and between SRH and mild diseases among women only (Benyamini et al., 2000), further support the perspective that women consider a wider range of health dimensions than men when processing information for the assessment of general health. Also, findings that men with a recent history of transient ischemic attack, but not community-dwelling men, have more accurate self-reports of their global health than their female counterparts suggest that previous encounters with a health problem may have a differential impact on men's and women's reporting behaviors (Dave et al., 2013; Dey et al., 2015).

Other studies challenge gender stereotypes in the context of seeking medical advice and reporting health problems. Analysis of primary healthcare data in the UK revealed very small gender differences in the number of doctor consultations within 24 months prior to the diagnosis of three cancers (Wang et al., 2014). Similarly, another study found little support for the hypothesis that women are more likely to consult professionals for back pain than men with similar morbidity, although there was some evidence for more active help-seeking behavior for headache in women than in men (Hunt et al., 2011). A recent qualitative study also found that men and women who were interviewed after receiving a lung cancer diagnosis had similar interpretations of and reactions to symptoms of the disease (MacLean et al., 2017). Contradicting the prevailing assumption of over-reporting among women, MacIntyre and colleagues (1999) found no evidence of gender differences in the reporting of health problems, irrespective of their seriousness and type. Additionally, research showing no gender differences in self-rated health at the same level of morbidity and in its predictive ability for hospitalizations runs counter to the hypothesis that women and men use different standards for assessing self-reported health (Case and Paxson, 2005). To directly tackle the question of whether women over-report and men under-report health conditions, researchers have compared self-reports to clinical data and found mixed results. Some suggest higher accuracy of self-reports among women than among their male counterparts (Dave et al., 2013), while others indicate the opposite (Dey et al., 2015; Short et al., 2009). However, we know relatively little about gender differences in the reporting of SRH.

SRH is the most frequently used indicator of health in social, economic, and epidemiological research, as it is a strong predictor of mortality (Mossey and Shapiro, 1982) and easy to include in surveys. Previous work on the meaning of SRH suggests that one of its main determinants across various populations is physical health (Au and Johnston, 2014; Hardy et al., 2014; Idler and Benyamini, 1997; Manderbacka, 1998). Jylhä (2009) developed a conceptual model for SRH to enhance researchers' understanding of the different factors that may influence an individual's perception of his or her health. According to this model, the evaluation of own health includes the review of information about biological health—e.g., medical diagnoses, functional status, symptoms and signs of illness—as well as lifestyle behaviors, and this evaluation is influenced by contextual social and cultural factors. Among these factors are the use of various reference groups (e.g., peers vs. younger/older persons), earlier experiences (e.g., experiencing pain or being diagnosed with a condition), and cross-cultural differences in using scales (e.g., linguistic differences in response options) (Jylhä, 2009). If contextual differences in the evaluation of SRH across cultures or socio-demographic groups are large, comparisons of SRH may lead to misleading results, as the observed differences will reflect not only the variations of true health across these groups but also the differences in reporting behavior.

Various approaches have been used to account for reporting heterogeneity and to improve the comparability of self-reported health measures across socio-demographic and cultural groups, including by gender and educational attainment (Jürges, 2007; Layes et al., 2012; Salomon, 2004; Schneider et al., 2012). One commonly used method is anchoring vignettes, which are brief texts describing a hypothetical situation (e.g., the level of health) which respondents are asked to evaluate using the same ordinal scale as for their own self-ratings (King et al., 2004; Salomon, 2004). Because the vignette is fixed for all respondents, variation in ratings is interpreted as a measure of reporting heterogeneity. Using anchoring vignettes, Grol-Prokopczyk et al. (2011) showed that in the Wisconsin Longitudinal Study the female advantage in SRH disappeared after accounting for reporting heterogeneity. However, in the Health and Retirement Study (HRS) men had more optimistic reporting in the domains of sleep, mobility, shortness of breath, and depression, and more pessimistic assessments for pain and memory compared with women (Dowd and Todd, 2011). Anchoring vignettes have also been used to examine differences in reporting behavior by education, wealth, and race (Bago d'Uva, O'Donnell, & Van Doorslaer, 2008; Dowd and Todd, 2011; Rossouw, Bago d'Uva and van Doorslaer, 2018). In the Survey of Health, Ageing and Retirement in Europe (SHARE), the direction and significance of educational inequalities before and after accounting for reporting heterogeneity based on vignette responses varied across six selected health domains –pain, sleep, mobility, emotional health, cognition, and breathing – and across countries (Bago d'Uva et al., 2008). For example, although Dutch persons with high education tended to assess their health more critically compared to their lower educated peers for most domains except cognition, this pattern was not evident among Swedes. Educational differences in pain and memory domains remained almost unchanged before and after adjustment for reporting heterogeneity in the HRS sample as well, but increased substantially for shortness of breath, depression, and mobility after the adjustment (Dowd and Todd, 2011).

While these findings are compelling, two important assumptions are made when using anchoring vignettes: response consistency and vignette equivalence (King et al., 2004; Salomon, 2004). Response consistency implies that an individual evaluates both specific health questions and related hypothetical scenarios in the same way, while vignette equivalence requires that the underlying health level depicted in each vignette be understood in the same way by all respondents, independent of socio-demographic or other characteristics. Although earlier studies found no major violations of response consistency and vignette equivalence (Bago d'Uva et al., 2008; Grol-Prokopczyk et al., 2011), recent studies, which have used stricter statistical methods to test these two assumptions, provide clear evidence that respondents from different cultures and socio-demographic groups perceive vignette texts as depicting fundamentally different levels of health (Bago d'Uva, Lindeboom, O'Donnell and van Doorslaer, 2011; Grol-Prokopczyk et al., 2015). These findings suggest that responses to the health vignettes in these studies cannot be used to correct for heterogeneity in health reporting.

In another approach, Layes et al. (2012) created a preference-standardized health-related quality of life measure designed to represent respondents' latent, true health in a sample of Canadian men and women. They interpreted systematic deviations between this measure and SRH as reporting behaviors. The authors found that men assessed their health to be significantly lower and women reported their health to be significantly higher relative to the “average Canadian”, but the magnitude of these gender differences was very small.

An alternative and arguably more reliable method to account for heterogeneity in health reporting is to use (semi-)objective information about health to adjust self-reported data (Jürges, 2007; Schneider et al., 2012). Using this approach, Jürges (2007) and Rebelo and Pereira (2014) showed that based on self-reports the healthiest individuals in Europe reside in Denmark and Sweden, whereas the least healthy live in Italy and Spain. However, when cross-cultural differences in reporting behavior were accounted for, these cross-national variations in general health were reduced, and the order of the countries from most to least healthy changed substantially.

Adapting this method, which we describe in detail below, the present study investigates the extent to which adjusting for differences in reporting behavior modifies gender differences in health and whether these changes are due to men and women over- and/or under-reporting their health. Since most previous studies have either focused on domain-specific measures of health or considered only positive evaluations of health, we also examine whether gender differences in reporting patterns are similar for the two opposite evaluations of SRH health: poor and good. Following Jylhä’s (2009) conceptual model, we separate SRH into two components: true health and the influences of contextual characteristics on the evaluation of health, i.e. reporting behavior. Since we assume that women and men will differ in both their true health and reporting behavior, we expect to find that an initial male advantage in SRH lessens or even disappears after adjusting for reporting behaviors. If our analyses show either that women over-report and men under-report poor health, or that women under-report while men over-report good health, stereotypical expectations about gender differences in health reporting will be supported. We may thus observe a reduction of the gender differences in both poor and good self-reported health when accounting for these gendered reporting behaviors. On the other hand, if women under-report and men over-report poor health, or if women over-report and men under-report good health, this would challenge stereotypes about gendered reporting behaviors. Finally, it is possible that women and men do not differ much in their reporting behaviors, and that the gender gaps in health will remain unchanged after adjusting for them. We might also expect that the relationship between reporting behavior and education differs among men and women. As reviewed above, although analyses of HRS data revealed quite comparable reporting patterns by education in the two genders, German data has suggested that reporting heterogeneity between men and women is driven in part by socioeconomic factors (Dowd and Todd, 2011; Schneider et al., 2012). Therefore, we also explore whether the observed reporting behaviors of men and women are consistent among individuals with different levels of education.

2. Materials and methods

2.1. Data and study population

We used data collected during Wave 1 (2004) of the Survey of Health, Ageing and Retirement in Europe (SHARE), which consists of 11 EU countries – Austria, Belgium, Denmark, France, Germany, Greece, Italy, Netherlands, Spain, Sweden, and Switzerland – and Israel (Börsch-Supan et al., 2013; Stuck et al., 2018). The SHARE is a multidisciplinary and cross-national panel survey of community-dwelling individuals aged 50 and older. The initial total sample of 29,373 individuals was reduced by 6.9 percent to 27,345 individuals (54% women) as a result of excluding persons with missing data on SRH (n = 2028) and/or independent variables. The major sources of missing data on independent variables were grip strength (n = 1707) and/or depression (n = 468). However, the sex-specific distributions of SRH were similar between those who had complete and missing data.

Designed to provide a comprehensive assessment of aging adults' well-being, the SHARE survey contains extensive questions about respondents’ health, making it an ideal source of data for this study. The key global health question asked interviewees to evaluate their health in general (“Would you say your health is … ?”) with 5 possible responses: very good, good, fair, bad, very bad (the EU version of the responses). Other health measures included in our analysis are physician-diagnosed reported chronic conditions, Basic Activities of Daily Living (BADL), mobility limitations, Instrumental Activities of Daily Living (IADL), physical performance tests (handgrip strength and walking speed), mental health, anthropometric measurements, and lifestyle behaviors. We used these characteristics to estimate a continuous latent health measure in the process described next.

2.2. Statistical approach

In this subsection we provide a more conceptual discussion of the statistical approach, while technical details about our methodology are given in the Appendix. To summarize, we employ a method used by Jürges (2007) to explore cross-national differences in reporting general health, modifying it to assess gender differences in health reporting. According to this approach, when responding to a survey question about their general health, participants assess their true health, which is measured on a continuous scale and is unobserved, and translate it onto a provided discrete 5-point scale. The thresholds that each individual uses to categorize their true health into a specific response option may be affected by the choice of a reference group, earlier health experiences, and cross-cultural differences in using scales, and thus, may differ across individuals depending on their gender, age, cultural background, education, and personality traits, among other factors. Following Jürges’ (2007) methodology, we computed a continuous estimate of individuals' underlying, latent health based on a wide range of health measures included in the data, and accounting for variations in reporting across socio-demographic and cultural groups. We then recoded these continuous health latent estimates into five categories that reflect the five response options on SRH. The result is a categorical health measure that is comparable to SRH, but which adjusts for heterogeneity in individuals’ reporting styles.

Providing some more specifics about this method, we first fitted a generalized ordered probit model that regresses the original SRH measure on two sets of independent variables (W. H. Greene and Hensher, 2010; W. Greene, Harris, Hollingsworth and Weterings, 2014; King et al., 2004). This model was proposed by King et al. (2004), and is an extension of a standard model for ordinal dependent variables. In contrast to the standard ordered probit models, the generalized ordered probit model relaxes the assumption that individuals use a common scale when rating their own health and allows for distinguishing between health and reporting differences. The first set of variables (health variables) included in the regression model assesses various specific aspects of individuals' health, including chronic conditions, mobility level, difficulties with daily activities, performances on grip strength tests, anthropometric measures, and lifestyle behaviors (Table 1). Using the second set of independent variables (threshold variables), the model identifies cut points between adjacent SRH response categories as functions of individual characteristics, i.e., it adjusts for socio-demographic and cultural differences in how the continuous latent health is projected onto the five-category SRH measure. These threshold variables included gender, age group, education, and country. By incorporating both health and threshold variables into the model, we obtained a measure of individuals’ continuous latent health (h_i, predicted latent health values) that controls for reporting heterogeneity.

Next, the model-derived predicted latent health values (h_i) were used to calculate disability weights for each health variable. These provide information about the individual impact of the specific health measures listed in Table 1 on the latent health construct, i.e., the extent

Table 1
Distribution of sociodemographic and health characteristics and estimated disability weights in the study population by gender.

	Women (n = 14,680)	Men (n = 12,663)	Total (n = 27,345)	Disability weight
Health variables (%)				
Parkinson's disease	0.4	0.6	0.5	0.152
Poor mobility	27.3	14.3	21.3	0.114
Grip test unable	2.7	1.9	2.3	0.101
Respiratory diseases	8.0	8.7	8.3	0.094
Heart attack/ other hearts problems	9.4	15.7	12.4	0.091
Stroke/other cerebral vascular disease	3.0	3.9	3.5	0.088
Cancer	5.9	4.8	5.4	0.081
Other diseases	17.9	15.7	16.9	0.080
ADL	9.8	7.7	8.8	0.079
Diabetes	10.1	11.4	10.7	0.067
Musculoskeletal diseases	31.3	14.0	23.3	0.064
Depression	46.2	27.4	37.5	0.062
Hypertension	37.5	34.5	36.1	0.051
Low grip strength	35.4	35	35.2	0.047
IADL	19	10.9	15.3	0.046
Underweight	1.8	0.4	1.1	0.041
Stomach, duodenal or peptic ulcer	5.3	6.8	6.0	0.036
Hip or femoral fracture	2.1	1.7	1.9	0.027
Chronic diseases 2+	44.4	38.7	41.7	0.027
Current smoker	15.5	23.5	19.2	0.026
Obese	18.2	16.2	17.3	0.019
High blood cholesterol	24.9	25.5	25.2	0.018
Overweight	36.6	49.9	42.8	0.015
Threshold variables (%)				
Education				
Low	56.1	45.5	51.2	
Secondary+	43.9	54.5	48.8	
Age groups				
50–59	37.7	36.5	37.1	
60–69	31.3	33.5	32.3	
70–79	21.4	22.2	21.8	
80+	9.7	7.8	8.8	
Country				
Belgium	12.5	13.1	12.8	
Sweden	10.0	10.6	10.3	
Germany	9.9	10.2	10.1	
France	10.3	9.6	10.0	
Netherlands	9.7	10.1	9.9	
Greece	8.7	9.0	8.8	
Italy	8.6	8.3	8.4	
Spain	8.4	7.2	7.8	
Israel	7.5	7.7	7.6	
Denmark	5.6	5.8	5.7	
Austria	5.4	4.8	5.1	
Switzerland	3.3	3.5	3.4	

to which latent health is reduced by the presence of Parkinson's disease, poor mobility, and other specific health measures. The disability weight for each health variable is a standardized coefficient of the from the generalized ordered probit model, and is equal to the ratio of the corresponding health coefficient to the difference between the lowest and highest values of predicted latent health (equation (3) in the Appendix). The predicted latent health values were also used to create an individual health index (H_i). The H_i is a proxy for true underlying health, and varies from 0 representing the (model-based) worst health state to 1 representing the (model-based) best health in the sample (equation (2) in the Appendix). H_i is, thus, a standardized predicted health value for an individual i in our study population. Disability weights reduce H_i by some given amount or percentage computed for an average individual in the study population (i.e., the H_i of each individual is reduced by the same amount if a heart attack or another heart problem is present).

Finally, H_i was reclassified into a five-category health measure that is adjusted for inter-individual differences in reporting behavior. To do so, we followed the original distribution of SRH categories for the entire study population irrespective of other characteristics (Jürges, 2007; Rebelo and Pereira, 2014). That is, if 10% of the study participants assessed their health to be very bad, the value of H_i, which corresponded to the lowest 10% of the H_i distribution was used as the cutoff level between the lowest two adjacent categories, i.e., "very bad" and "bad" (see Supplementary Table 1 and Supplementary Fig. 1). Although the original response frequencies were maintained, because we adjusted for heterogeneity in reporting behavior between men and women and across age groups and countries, a study participant who initially perceived her/his health as very bad may actually have been reclassified as having fair health. In the end we obtain a categorical measure of health similar to SRH, but which adjusts for inter-individual reporting heterogeneity—or in other words, that takes into account how different people transform their "real" latent health into the ordinal SRH categories.

All health determinants that comprised the first set of variables used to predict the latent health measure were coded as binary. Some variables identify the presence of the following physician-diagnosed reported chronic conditions ("Has a doctor ever told you that you had any of the conditions on this card?"): 1) heart attack or other heart problems, 2) hypertension or use of anti-hypertensive medications, 3) high blood cholesterol or use of statins, 4) stroke or other cerebrovascular diseases, 5) diabetes or use of antidiabetic medications, 6) chronic respiratory diseases including asthma, 7) musculoskeletal diseases, 8) cancer (including leukemia or lymphoma, but excluding minor skin cancers), 9) stomach, duodenal, or peptic ulcer, 10) Parkinson disease, 11) hip/femoral fracture, 12) other chronic conditions, and 13) the presence of two or more chronic diseases as a quasi-interaction. Other determinants reflected participants' mobility level and difficulties with a range of activities: 14) poor mobility (three or more reported mobility limitations, i.e. the ability to engage in activities broadly ranging from sitting for about 2 h to climbing several flights of stairs without resting), 15) Activities of Daily Living (ADL) disability (one or more reported ADL limitations, i.e. the ability to perform such simple tasks as dressing, including putting on shoes and socks or walking across a room), and 16) Instrumental Activities of Daily Living (IADL) disability (one or more reported IADL limitations, i.e. the ability to manage tasks necessary for fully independent life at own home). We also included 17) low grip strength (lowest sex-specific tertile (< 23 kg for women and < 39 kg for men) and 18) unable to perform grip strength test as objective measures of health. The three anthropometric measures based on self-reports in the analysis are 19) underweight (BMI ≤ 20), 20) overweight (BMI > 20 and BMI < 25), 21) obese (BMI ≥ 30). Finally, we accounted for mental health in identifying those with 22) a score of 4 and above on the EURO-D depressive symptoms scale, and a lifestyle behavior with 23) being a current smoker. Threshold variables include gender, education level (primary, i.e. the International Standard Classification of Education (ISCED) level 2 or less, vs. secondary+, i.e. ISCED levels 3 and above), country (12 countries), and age groups (from 50 to 80 + in 10-year intervals).

Although in the initial model we included all health variables suggested by Jürges (2007) and Rebelo and Pereira (2014), we limited the final model to those significantly related to respondents' original SRH. While walking speed was significant in the initial model, we excluded it from the final model due to the large number of individuals with missing values. We also combined arthritis, including osteoarthritis and rheumatism, and osteoporosis into musculoskeletal problems, while asthma and chronic lung disease were combined into respiratory problems. We performed all analyses with a 3-level measure of education, i.e., low (ISCED level 2 or less), medium (ISCED levels 3 and 4), and high (ISCED level 5 or higher). Since the results were very similar in medium and high education groups, we opted for a more parsimonious models with 2-level education.

All analyses were performed in R using the *hopit* package (Dańko, 2019). A full description of the method and calculations applied in this study is given in Appendix 1. The package is now available online: https://github.com/MaciejDanko/hopit. The introduction to the package can be found here: https://github.com/MaciejDanko/hopit/blob/master/vignettes/introduction_to_hopit.pdf.

3. Results

Table 1 shows the distribution of health variables (health-related characteristics) and threshold variables (socio-demographic characteristics) of the study population by gender, as well as the estimated disability weights for each health variable. The prevalences of mobility limitations, musculoskeletal problems, depression, IADL disabilities, and underweight were substantially greater among women, while having had a heart attack/another heart problem, stroke/other cerebrovascular disease, ulcer, and being a current smoker or overweight was more likely for men. In terms of disability weights, Parkinson's disease, poor mobility, being unable to perform the grip strength test, and having a history of heart attack and stroke had the largest contribution to the reduction of health. The lowest contributions were from high total cholesterol, being overweight or obese, and being a current smoker. Women were more likely to be in the lower education group than men, while the age distribution was similar across the two genders.

Although all models were fitted with 5-level SRH as a dependent variable, to facilitate the interpretation of the results we combined the response options very bad and bad (hereafter poor health) and the response options very good and good (hereafter good health). To determine whether the original and adjusted prevalences of poor/good health are different, we examined whether the original prevalences are within 95% confidence intervals (CIs) of the adjusted ones.

3.1. Self-reported vs. adjusted poor health by age and gender

The prevalence of poor SRH on the original scale increased with increasing age from 5.9% among 50–59 year old women to 20.4% among women aged 80+, and from 5.7% among 50–59 year old men to 15.5% among men age 80+ (Fig. 1).

When SRH was adjusted for differences in reporting styles, the proportion of women with poor health declined slightly in the youngest age group (age 50–59) and increased among older women (Fig. 1). Similarly, among men the prevalence of poor health was lower for the two younger age groups (age 50–59 and 60–69) than the original estimates, remained almost unchanged among men 70–79 years old, and increased substantially among men age 80+.

Gender differences (the absolute difference between the prevalence of poor health among men and women) in poor health on the original scale were small and not significant among persons in their 50s (0.21%) and 60s, and they increased with increasing age, with a difference of 4.9% in the oldest age group (Fig. 2, left panel). After adjusting for reporting heterogeneity, the proportion of those with poor health was substantially higher among women than among men at all ages, and gender differences in poor health widened across all age groups (Fig. 2, left panel). As with the original scale, gender differences in the adjusted fractions increased with advancing age (from 2.0% among 50–59 year olds to 9.6% in the oldest age group).

To help interpret whether these widened gender differences in poor health were due to men and women under- and/or over-reporting their poor health, we plotted the differences between the adjusted and original proportions of men and women in poor health by age. Negative differences in the left panel of Fig. 3 mean that the adjusted prevalences of poor health are smaller than the original prevalences, indicating that persons report lower levels of health than they have (i.e., they tend to over-report poor health). In contrast, positive differences between the adjusted and original prevalences of poor health suggest that persons report better health than they have (i.e., they tend to under-report poor health). Fig. 3 (left panel) shows that 50-59 year-old women and 50–59 and 60-69 year-old men reported worse health than they actually had. The difference between adjusted and original prevalences of poor health among 70-79 year-old men was negligible, suggesting that their reporting of poor health was fairly accurate. By contrast, women age 60 + and men age 80 + reported better health than they have. These results also imply that, although the male advantage in general health increases after adjusting for reporting heterogeneity, there are no clear gender-specific patterns in reporting behaviors. Rather, the reporting of poor health appear to be age-dependent, with both younger men and women over-reporting poor health and older individuals under-reporting poor health.

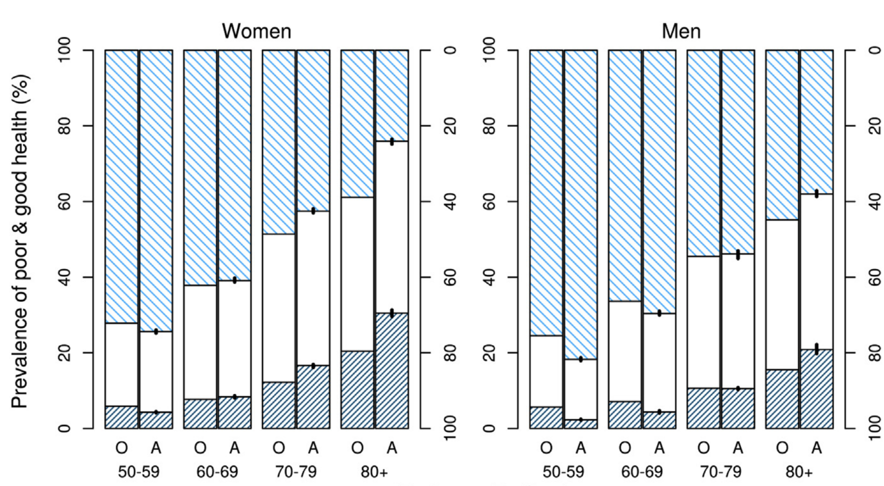

Fig. 1. Original (O) and adjusted (A) prevalences of poor and good health among women and men by age.

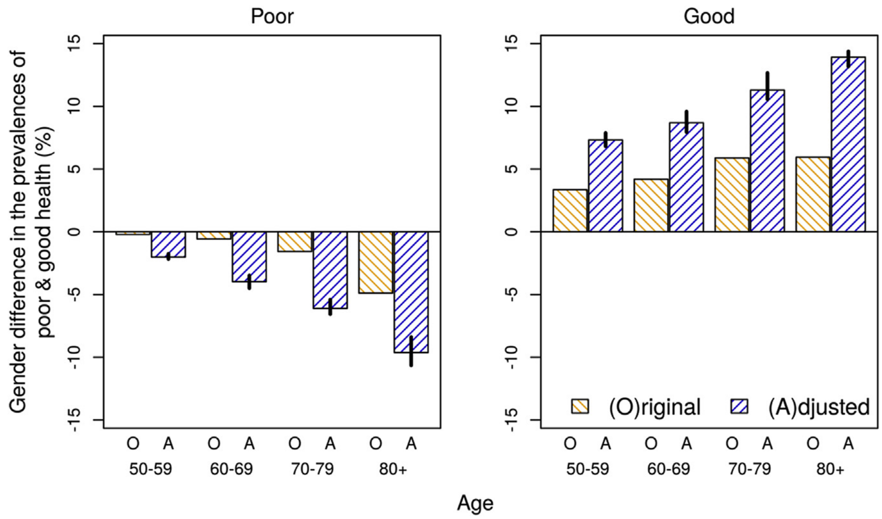

Fig. 2. Gender (male – female) differences in the prevalences of poor and good health by age.

3.2. Self-reported vs. adjusted good health by age and gender

The percentage of persons with good SRH declined with increasing age from 72.1% among 50-59 year-old women to 38.9% among women aged 80 + , and from 75.5% among 50-59 year-old men to 44.9% among the oldest men (Fig. 1). Relative to the original proportions, the adjusted prevalences of good health increased slightly among youngest women (50–59 years) and declined in all other age groups. When reporting heterogeneity was adjusted for, the prevalence of good health increased among men aged 50 to 69, remained almost unchanged among men aged 70–79, and declined in the oldest age group.

The percentage of people reporting good health on the original scale was higher among men than in women at all ages. Gender differences increased from 3.6% among persons aged 50–59 to 5.9% in the oldest age group and were statistically significant at all ages (Fig. 2, right panel). After controlling for reporting behaviors, the gender gaps in the proportions of good health favoring men widened at all ages, and the magnitude of gender differences increased with increasing age (Fig. 2, right panel) from 7.3% in the youngest age group to 13.9% in the oldest age group.

In Fig. 3 (right panel), a positive difference between the adjusted and original prevalences of good health suggests that fewer persons reported being in good health than they actually had, i.e. they under-reported their good health. A negative difference the adjusted and original prevalences of good health implies that more people reported having good health than they actually had, i.e. they over-reported their good health. Thus, Fig. 3 (right panel) shows that the youngest women under-reported good health, while more older women over-reported their good health.

Fig. 3 (right panel) also indicates that younger men (50–59 and 60–69 years) under-reported being in good health, while men age 70–79 reported being in good health fairly accurately. However, as among women, more men at the oldest ages over-reported their good health. In short, these findings suggest that in our sample the reporting pattern among the youngest and oldest men is similar to that of their female peers. Among men and women in their 60s and 70s, the reporting behaviors are more mixed with no clear gender-specific patterns.

3.3. Self-reported vs. adjusted poor and good health by education

We also examined whether changes in the initial vs. adjusted proportions of poor/good health vary by education, as well as whether accounting for educational differences affects gender gaps in SRH. Although low educated women tended to have higher prevalence of poorer health than their male counterparts, the gender differences were not statistically significant with exception of the oldest age group (Supplementary Fig. 2). The female disadvantage in poor health among

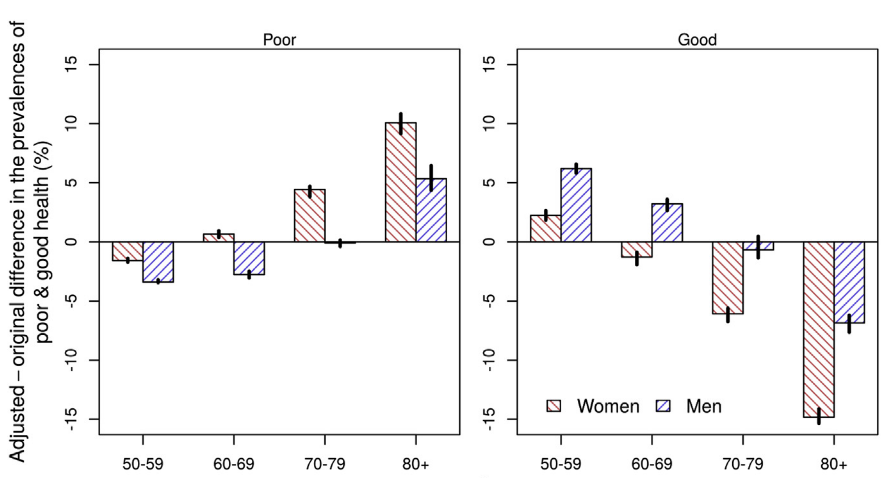

Fig. 3. Differences between the adjusted and original prevalences of poor and good health by gender and age.

the persons with secondary + education group varied across age groups. When differences in reporting behaviors were adjusted for, the female disadvantage in poor health became consistent across all ages and in both education groups, and the magnitude of gender gaps in the prevalence of poor health was greater in low education groups than among better educated groups.

The differences between the adjusted and original proportions of men and women with poor health by education suggest that the patterns of health reporting among men and women with low and secondary + education were similar, and they were similar to the general patterns indicated earlier (see section *Self-reported vs. adjusted poor health by age and gender*) (Supplementary Fig. 3). Our findings show that on the original scale the female disadvantage was more pronounced in the low education group than among better educated groups, while it was not always apparent among individuals with secondary + education. After adjusting for reporting heterogeneity, the female disadvantage in poor health widened in low education groups and was observed among better educated persons across all ages. Irrespective of education level, younger women (50–59 years) and men (50–69 years) over-report poor health, while women aged 60 + and men aged 80+ under-report poor health. Men at ages 70–79 report having poor health relatively accurately.

All of the general patterns observed earlier for good health were similar in the two education groups. Our analyses indicate that regardless of education level, younger men and women tend to under-report good health, while older individuals of both genders tend to over-report good health (Supplementary Fig. 3). Our findings of no consistent changes in the prevalences of poor and good health after adjusting for reporting heterogeneity in low vs. secondary + education suggest that the effect of education on reporting behavior is similar in the two genders.

3.4. Sensitivity analysis

Although it is well established that chronic conditions are important determinants of SRH (Jylhä et al., 1986; Singh-Manoux et al., 2006), less research has investigated whether the effects of chronic conditions on SRH differ when considered as single vs. as multimorbid conditions (Mavaddat et al., 2014). Mavaddat et al. (2014) revealed that the odds of reporting poor health increased with the number of chronic conditions, and this association was stronger for men than for women. To test whether including physician-diagnosed reported chronic conditions as multimorbid rather than single conditions has an impact on the estimation of disability weights and adjusted proportions of poor and good health, we performed additional analyses fitting a model where specific chronic conditions were replaced with a categorical variable indicating the number of chronic diseases from 0 to 9+. All gender- and age-specific patterns were similar to those found for the original models (results available on request).

Another potential concern is related to differences in SRH created by variations in the provided response options. The SHARE questionnaire also included a version of the general health question with response options commonly used in the U.S. – excellent, very good, good, fair, poor. We replicated our analysis with this alternative version of response options. Although some small differences were observed, the overall results by gender and education levels were very similar, and our conclusions remained the same (results available on request).

4. Discussion

Although previous works have suggested that the female disadvantage in SRH can be partially explained by women's greater attention to bodily symptoms and their willingness to report health problems, research rigorously testing this assumption is limited. This study examined whether adjusting for differences in reporting behaviors modifies gender differences in health, and whether these changes are due to men and women over- and/or under-reporting their health.

We found small gender differences in the prevalence of poor health on the original scale among persons in their fifties and sixties, and a consistent male advantage in the prevalence of both poor and good health among older persons. After adjusting for differences in men's and women's reporting behaviors, gender differences in both poor and good health widened. Contrary to widespread assumptions about gender-stereotypical reporting behaviors which suggest that women over-report poor health and men over-report good health, we found no clear evidence for gender-specific patterns in reporting of either poor or good health. Rather, health reporting varied greatly by age: younger (50–59 and 60–69 year old) women and men in our study population tended to over-report poor health and under-report good health, while the oldest women and men tended to under-report poor health and over-report good health. Men in their seventies had fairly accurate reporting of both poor and good health. Although gender differences in poor and good health became more apparent in both education groups after adjusting for reporting behavior, there was no consistent pattern in over- or under-reporting of health across education groups and in the magnitude of the female disadvantage in health. These findings suggest that reporting behaviors were similar among men and women with lower and secondary + education.

Our results are consistent with a Canadian study showing that being 80 years of age and older was the strongest determinant of an optimistic health evaluation, i.e., above the average Canadian (Layes et al., 2012). Extensive literature on adaptation to negative conditions suggests that coping strategies and lower expectations in old age protects individuals from the loss of life satisfaction even when experiencing somatic problems and socioeconomic adversities (Ebner et al., 2006; Jopp and Rott, 2006; Staudinger et al., 1999). Individuals seem to adapt emotionally to their constant pain, discomfort, and disabilities caused by chronic conditions, which may underlie a shift toward reporting better subjective health over the disease trajectory, even if true health is not improving (Damschroder et al., 2005). In a longitudinal analysis of the British Cohort Study, Cubí-Mollá et al. showed that although suffering from a long-term illness deceased the probability of reporting excellent health, the length of illness was positively related to reporting better health (Cubí-Mollá et al., 2017). Other scholars revealed that despite substantial declines in physical and mental health with increasing age, middle-aged and young-old Danes who participated in three longitudinal population-based surveys were able to maintain their initial levels of life satisfaction to very old ages (Vestergaard et al., 2015).

Our findings of no clear gender patterns in reporting behavior partially agree with previous research using the Wisconsin Longitudinal Study (Grol-Prokopczyk et al., 2011) and the Health and Retirement Survey (Dowd and Todd, 2011), which revealed mixed patterns for gender differences in health depending on the selected health domain. While men tended to provide relatively optimistic evaluations of their health, adjusting for reporting heterogeneity had no substantive impact on gender differences in mobility, diminished female disadvantage in depression and sleep, and increased male disadvantage in shortness of breath (Dowd and Todd, 2011). Although these studies used domain-specific rather than general measures of health, like our findings, their results are not always in line with dominating assumptions that women tend to over-report and men tend to “hide” their health problems. Interestingly, a Canadian study found the opposite patterns: men evaluated their health more pessimistically (lower relative to the average Canadian) and women assessed their health more optimistically, but these gender differences were very small (Layes et al., 2012).

In line with prior work, we find that education inequalities in health widen after adjusting for reporting heterogeneity. However, our findings indicate that reporting patterns do not systematically differ among persons with lower and secondary + education and that gender gaps in health are similar across the two education groups. Although a previous study based on the SHARE data showed that correcting for these differences generally widened educational inequalities in health, this enlargement was attributable to more negative ratings of a given health state by better educated individuals (Bago d'Uva et al., 2008). Possible explanations for these inconsistent findings may underlie in the methodological differences between this study and our own our, i.e. the use of anchoring vignettes vs. other (semi-)objective health measures to adjust for reporting heterogeneity and domain-specific measures of health vs. global health.

Previous work regarding educational differences in reporting behaviors is contradictory. A study based on a Canadian community sample showed that higher socioeconomic groups assessed their health more pessimistically than their peers from low socioeconomic groups, but the differences by education were very small albeit statistically significant (Layes et al., 2012). In other North American-based studies, the accuracy of reporting of cardiovascular risk factors, such as hypertension or hypercholesterolemia, was similar among people with different levels of educational attainment (Dey et al., 2015; Vargas et al., 1997). There is, however, some evidence that the concordance between self-reports and administrative data regarding health care utilization and absenteeism was higher among persons with an advanced educational degree than those with less education (Short et al., 2009).

A strength of the present study is that the measure of health considered takes into account the multi-dimensionality of global health by using a wide range of health-related characteristics to better get at the level of "true" health. However, a substantial part of the health characteristics used in the present study are self-reports, and research evidence regarding the accuracy of self-reported chronic conditions across socio-demographic groups is conflicting. Some studies found that men were more likely to accurately report hypertension, hypercholesterolemia, and diabetes (Dey et al., 2015), as well as the number of doctor visits, absenteeism, and cardiovascular risk factors than women (Short et al., 2009). Others revealed higher odds of over-reporting of chronic diseases or even greater validity of self-reported hypertension among women than among men (Dave et al., 2013; Galenkamp et al., 2014). Some of these discrepant findings can be attributable to differences in the study populations, i.e., patient vs. community-based samples, and/or previous experiences of health problems, and to the intensity of healthcare use (Bhandari and Wagner, 2006; Oksuzyan et al., 2009).

Further, some individuals may be unaware of a serious health problem they are suffering from, e.g. silent hypertension, myocardial infraction, and ischemic stroke, which could bias our estimates of disability weights. Although the prevalence of these silent chronic conditions varies across study populations depending on the patients' ages and the method used to detect the condition, in the general population the prevalence appears to be rather small (up to 5%) with no clear sex-specific pattern (Das et al., 2008; Valensi et al., 2011). Another study limitation may underlie in the methodology we applied to adjust for reporting heterogeneity. Despite being comprehensive, this method may not be able to eliminate entirely the heterogeneity in reporting behavior as the set of threshold variables may be incomplete. It does not include, for example, profession/occupation, partnership status, religion, personality characteristics and other features that may potentially influence reporting of health. Additionally, some characteristics that are used as threshold variables may not only modify individuals' reporting behavior, but also influence their health. Jürges's model (2007) was further developed by Rebelo and Pereira (2014) to allow a variable to be included as both a threshold and a health variable. However, their findings showed that most of these variables were not statistically significantly related to SRH as health variables, but only as threshold variables. Since we limited our model to those health variables which were significantly associated with SRH, we decided to use country and education only as threshold variables. Further, the disability weights were computed for an average individual in our study population, implying that the presence of a specific health condition reduces estimated latent health to the same extent for all individuals. Considering previous research reports on sex-specific associations of body mass index, waist-to-hip ratio, and electrocardiographic indicators with SRH, the associations between health characteristics and SRH in our generalized order probit model and, thus, the extent to which disability weights reduce perfect health, may also be gender-specific (Oksuzyan et al., 2015). Considering the above-mentioned limitations, our future studies will focus on investigating how adjusting for reporting behavior modifies gender differences in general health using other survey datasets where biomarker data are available. Also, since gender-related social norms may vary across European countries, it is possible that gender differences in reporting behavior differ across the SHARE countries, which is a topic for future investigation.

The present study adds to the existing literature on gender differences in reporting behavior by assessing whether adjusting for reporting styles modifies gender differences in SRH. Our findings challenge prevailing gender stereotypes that "sensitive" women over-report and "stoical" men under-report health problems. They also highlight the importance of attending to the health problems reported by women and men equally carefully, which is particularly salient for clinical settings and may help avoid delayed diagnosis and treatment of health problems more commonly seen in the opposite gender. Consolidating and extending similar observations in the U.K. (Hunt et al., 2011; MacLean et al., 2017), we argue that both research and medical communities should abandon traditional views that women more readily report poor health and men over-report good health, and that we should regard a malaise symptom, such as general weakness, as an equivalent sign of a slowly progressing potentially lethal disease in men and women. These steps and further in-depth research are needed to advance our understanding of gender differences in the experience and reporting of health and gender-specific barriers to the timely and appropriate use of healthcare services.

Acknowledgments

This work was supported by Max Planck Society within the framework of the project "On the edge of societies: New vulnerable populations, emerging challenges for social policies and future demands for social innovation. The experience of the Baltic Sea States (2016–2021)". Vladimir M. Shkolnikov was partly supported by the Russian program of the state support of leading universities of the Russian Federation "5–100".

This paper uses data from SHARE Wave 1 (DOIs: 10.6103/SHARE.w1.600), see Börsch-Supan et al. for methodological details (Börsch-Supan et al., 2013). The SHARE data collection has been primarily funded by the European Commission through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812) and FP7 (SHARE-PREP: N°211909, SHARE-LEAP: N°227822, SHARE M4: N°261982). Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C) and from various national funding sources is gratefully acknowledged (see www.share-project.org).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2019.03.002.

References

Au, N., Johnston, D.W., 2014. Self-assessed health: what does it mean and what does it hide? Soc. Sci. Med. 121, 21–28. http://doi.org/10.1016/j.socscimed.2014.10.007.

Bago d'Uva, T., Lindeboom, M., O'Donnell, O., van Doorslaer, E., 2011. Slipping anchor? Testing the vignettes approach to identification and correction of reporting

heterogeneity. J. Hum. Resour. 46 (4), 875–906. http://doi.org/10.2139/ssrn.1572319.
Bago d'Uva, T., O'Donnell, O., Van doorslaer, E., 2008. Differential health reporting by education level and its impact on the measurement of health inequalities among older Europeans. Int. J. Epidemiol. 37 (6), 1375–1383. http://doi.org/10.1093/ije/dyn146.
Bambra, C., Pope, D., Swami, V., Stanistreet, D., Roskam, A., Kunst, A., Scott-Samuel, A., 2008. Gender, health inequalities and welfare state regimes: a cross-national study of 13 European countries. J. Epidemiol. Community Health 63 (1), 38–44. http://doi.org/10.1136/jech.2007.070292.
Barsky, A.J., Peekna, H.M., Borus, J.F., 2001. Somatic symptom reporting in women and men. J. Gen. Intern. Med. 16 (4), 266–275. http://doi.org/https://doi.org/10.1046/j.1525-1497.2001.016004266.x.
Benyamini, Y., Leventhal, E.A., Leventhal, H., 2000. Gender differences in processing information for making self-assessments of health. Psychosom. Med. 62 (3), 354–364.
Bhandari, A., Wagner, T., 2006. Self-reported utilization of health care services: improving measurement and accuracy. Med. Care Res. Rev. 63 (2), 217–235. http://doi.org/10.1177/1077558705285298.
Bohannon, R.W., Peolsson, A., Massy-Westropp, N., Desrosiers, J., Bear-Lehman, J., 2006. Reference values for adult grip strength measured with a Jamar dynamometer: a descriptive meta-analysis. Physiotherapy 92 (1), 11–15. http://doi.org/10.1016/j.physio.2005.05.003.
Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., et al., 2013. Data resource profile: the survey of health, ageing and retirement in europe (share). Int. J. Epidemiol. 42 (4), 992–1001. http://doi.org/10.1093/ije/dyt088.
Brewster, K.L., Padavic, I., 2000. Change in gender-ideology, 1977-1996: the contributions of intracohort change and population turnover. J. Marriage Fam. 62 (2), 477–487. http://doi.org/10.1111/j.1741-3737.2000.00477.x.
Case, A., Paxson, C., 2005. Sex differences in morbidity and mortality. Demography 42 (2), 189–214. http://doi.org/10.2307/4147343.
Cooper, H., 2002. Investigating socio-economic explanations for gender and ethnic inequalities in health. Soc. Sci. Med. 54 (5), 693–706. http://doi.org/10.1016/S0277-9536(01)00118-6.
Courtenay, W.H., 2000. Constructions of masculinity and their influence on men's well-being: a theory of gender and health. Soc. Sci. Med. 50 (10), 1385–1401. http://doi.org/10.1016/S0277-9536(99)00390-1.
Crimmins, E.M., Kim, J.K., Solé-Auró, A., 2011. Gender differences in health: results from SHARE, ELSA and HRS. Eur. J. Public Health 21 (1), 81–91. http://doi.org/10.1093/eurpub/ckq022.
Cubí-Mollá, P., Jofre-Bonet, M., Serra-Sastre, V., 2017. Adaptation to health states: sick yet better off? Health Econ. 26 (12), 1826–1843. http://doi.org/10.1002/hec.3509.
Dahlin, J., Härkönen, J., 2013. Cross-national differences in the gender gap in subjective health in Europe: does country-level gender equality matter? Soc. Sci. Med. 98, 24–28. http://doi.org/10.1016/j.socscimed.2013.08.028.
Damschroder, L.J., Zikmund-Fisher, B.J., Ubel, P.A., 2005. The impact of considering adaptation in health state valuation. Soc. Sci. Med. 61 (2), 267–277. http://doi.org/10.1016/j.socscimed.2004.11.060.
Dańko, M.J., 2019. Hopit R-Package: Hierarchical Ordered Probit Models with Application to Reporting Styles. Retrieved from. https://github.com/MaciejDanko/hopit.
Das, R.R., Seshadri, S., Beiser, A.S., Kelly-Hayes, M., Au, R., Himali, J.J., et al., 2008. Prevalence and correlates of silent cerebral infarcts in the framingham offspring study. Stroke 39 (11), 2929–2935. http://doi.org/10.1161/STROKEAHA.108.516575.
Dave, G.J., Bibeau, D.L., Schulz, M.R., Aronson, R.E., Ivanov, L.L., Black, A., Spann, L., 2013. Predictors of congruency between self-reported hypertension status and measured blood pressure in the stroke belt. Journal of the American Society of Hypertension 7 (5), 370–378. http://doi.org/10.1016/j.jash.2013.04.007.
Dey, A.K., Alyass, A., Muir, R.T., Black, S.E., Swartz, R.H., Murray, B.J., Boulos, M.I., 2015. Validity of self-report of cardiovascular risk factors in a population at high risk for stroke. J. Stroke Cerebrovasc. Dis. 24 (12), 2860–2865. http://doi.org/10.1016/j.jstrokecerebrovasdis.2015.08.022.
Dowd, J.B., Todd, M., 2011. Does self-reported health bias the measurement of health inequalities in U . S . Adults ? Evidence using anchoring vignettes from the health and retirement study. J. Gerontol. 66, 478–489. http://doi.org/10.1093/geronb/gbr050.
Ebner, N.C., Freund, A.M., Baltes, P.B., 2006. Developmental changes in personal goal orientation from young to late adulthood: from striving for gains to maintenance and prevention of losses. Psychol. Aging 21 (4), 664–678. http://doi.org/10.1037/0882-7974.21.4.664.
Galenkamp, H., Huisman, M., Braam, A.W., Schellevis, F.G., Deeg, D.J.H., 2014. Disease prevalence based on older people's self-reports increased, but patient-general practitioner agreement remained stable, 1992-2009. J. Clin. Epidemiol. 67 (7), 773–780. http://doi.org/10.1016/j.jclinepi.2014.02.002.
Greene, W.H., Hensher, D.A., 2010. Modeling Ordered Choices. Cambridge University Press, Cambridge. http://doi.org/10.1017/CBO9780511845062.
Greene, W., Harris, M.N., Hollingsworth, B., Weterings, T.A., 2014. Heterogeneity in ordered choice models: a review with applications to self-assessed health. J. Econ. Surv. 28 (1), 109–133. http://doi.org/10.1111/joes.12002.
Grol-Prokopczyk, H., Freese, J., Hauser, R.M., 2011. Using anchoring vignettes to assess group differences in general self-rated health. J. Health Soc. Behav. 52 (2), 246–261. http://doi.org/10.1177/0022146510396713.
Grol-Prokopczyk, H., Verdes-Tennant, E., McEniry, M., Ispány, M., 2015. Promises and pitfalls of anchoring vignettes in health survey research. Demography 52 (5), 1703–1728. http://doi.org/10.1007/s13524-015-0422-1.
Hardy, M., Acciai, F., Reyes, A., 2014. How health conditions translate into self-ratings: a comparative study of older adults across europe. J. Health Soc. Behav. 55 (3), 320–341. http://doi.org/10.1177/0022146514541446.
Hunt, K., Adamson, J., Hewitt, C., Nazareth, I., 2011. Do women consult more than men? A review of gender and consultation for back pain and headache. J. Health Serv. Res. Policy 16 (2), 108–117. http://doi.org/10.1258/jhsrp.2010.009131.
Idler, E.L., Benyamini, Y., 1997. Self-rated health and Mortality : a review of twenty-seven community studies *. J. Health Soc. Behav. 38 (1), 21–37.
Jopp, D., Rott, C., 2006. Adaptation in very old age: exploring the role of resources, beliefs, and attitudes for centenarians' happiness. Psychol. Aging 21 (2), 266–280. http://doi.org/10.1037/0882-7974.21.2.266.
Jürges, H., 2007. True health vs response styles: exploring cross-country differences in self-reported health: TRUE HEALTH VS RESPONSE STYLES. Health Econ. 16 (2), 163–178. http://doi.org/10.1002/hec.1134.
Jylhä, M., 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://doi.org/10.1016/j.socscimed.2009.05.013.
Jylhä, M., Leskinen, E., Alanen, E., Leskinen, A.L., Heikkinen, E., 1986. Self-rated health and associated factors among men of different ages. J. Gerontol. 41 (6), 710–717.
King, G., Murray, C.J.L., Salomon, J.A., Tandon, A., 2004. Enhancing the validity and cross-cultural comparability of measurement in survey research. Am. Pol. Sci. Rev. 98 (1), 191–207. http://doi.org/10.1017/S0003055403000881.
Layes, A., Asada, Y., Kepart, G., 2012. Whiners and deniers - what does self-rated health measure? Soc. Sci. Med. 75 (1), 1–9. http://doi.org/10.1016/j.socscimed.2011.10.030.
MacIntyre, S., Ford, G., Hunt, K., 1999. Do women "over-report" morbidity? Men's and women's responses to structured prompting on a standard question on long standing illness. Soc. Sci. Med. 48 (1), 89–98. http://doi.org/10.1016/S0277-9536(98)00292-5.
Macintyre, S., Hunt, K., Sweeting, H., 1996. Gender differences in health: are things really as simple as they seem? Soc. Sci. Med. 42 (4), 617–624. http://doi.org/10.1016/0277-9536(95)00335-5.
MacLean, A., Hunt, K., Smith, S., Wyke, S., 2017. Does gender matter? An analysis of men's and women's accounts of responding to symptoms of lung cancer. Soc. Sci. Med. 191, 134–142. http://doi.org/10.1016/j.socscimed.2017.09.015.
Manderbacka, K., 1998. Examining what self-rated health question is understood to mean by respondents. Scand. J. Publ. Health 26 (2), 145–153. http://doi.org/10.1177/14034948980260020301.
Mavaddat, N., Valderas, J.M., Van Der Linde, R., Khaw, K.T., Kinmonth, A.L., 2014. Association of self-rated health with multimorbidity, chronic disease and psychosocial factors in a large middle-aged and older cohort from general practice: a cross-sectional study. BMC Fam. Pract. 15 (1), 1–11. http://doi.org/10.1186/s12875-014-0185-6.
Mossey, J.M., Shapiro, E., 1982. Self-rated health: a predictor of mortality among the elderly. Am. J. Public Health 72 (8), 800–808.
Oksuzyan, A., Crimmins, E., Saito, Y., O'Rand, A., Vaupel, J.W., Christensen, K., 2010. Cross-national comparison of sex differences in health and mortality in Denmark, Japan and the US. Eur. J. Epidemiol. 25 (7). http://doi.org/10.1007/s10654-010-9460-6.
Oksuzyan, A., Gumà, J., Doblhammer, G., 2018. Sex differences in health and survival. In: Doblhammer, J., Gabriele, Gumà (Eds.), A Demographic Perspective on Gender, Family and Health in Europe. Springer International Publishing, pp. 65–100.
Oksuzyan, A., Petersen, I., Stovring, H., Bingley, P., Vaupel, J.W., Christensen, K., 2009. The male-female health-survival paradox: a survey and register study of the impact of sex-specific selection and information bias. Ann. Epidemiol. 19 (7). http://doi.org/10.1016/j.annepidem.2009.03.014.
Oksuzyan, A., Shkolnikova, M., Vaupel, J.W., Christensen, K., Shkolnikov, V.M., 2015. Sex differences in biological markers of health in the study of Stress, Aging and Health in Russia. PLoS One 10 (6). http://doi.org/10.1371/journal.pone.0131691.
Palacios-Ceña, D., Jiménez-García, R., Hernández-Barrera, V., Alonso-Blanco, C., Carrasco-Garrido, P., Fernández-de-las-Peñas, C., 2012. Has the prevalence of disability increased over the past decade (2000-2007) in elderly people? A Spanish population-based survey. J. Am. Med. Dir. Assoc. 13 (2), 136–142. http://doi.org/10.1016/j.jamda.2010.05.007.
Rebelo, L.P., Pereira, N.S., 2014. Assessing health endowment, access and choice determinants: impact on retired europeans' (In)activity and quality of life. Soc. Indicat. Res. 119 (3), 1411–1446. http://doi.org/10.1007/s11205-013-0542-1.
Rossouw, L., Bago d'Uva, T., van Doorslaer, E., 2018. Poor health reporting? Using vignettes to recover the health gradient by wealth. Demography 1935–1956. http://doi.org/10.2139/ssrn.2932778.
Salk, R.H., Hyde, J.S., Abramson, L.Y., 2017. Psychological bulletin gender differences in depression in representative national samples: meta-analyses of diagnoses and symptoms gender differences in depression in representative national samples. Meta-Analyses of Diagnoses and Symptoms 143 (8), 783–822. http://doi.org/10.1037/bul0000102.
Salomon, J.A., 2004. Comparability of self rated health: cross sectional multi-country survey using anchoring vignettes. BMJ 328 (7434) 258–0. http://doi.org/10.1136/bmj.37963.691632.44.
Schneider, U., Pfarr, C., Schneider, B.S., Ulrich, V., 2012. I feel good! Gender differences and reporting heterogeneity in self-assessed health. Eur. J. Health Econ. 13 (3), 251–265. http://doi.org/10.1007/s10198-011-0301-7.
Short, M.E., Goetzel, R.Z., Pei, X., Tabrizi, M.J., Ozminkowski, R.J., Gibson, T.B., et al., 2009. How accurate are self-reports? Analysis of self-reported health care utilization and absence when compared with administrative data. J. Occup. Environ. Med. 51 (7), 786–796. http://doi.org/10.1097/JOM.0b013e3181a86671.
Singh-Manoux, A., Martikainen, P., Ferrie, J., Zins, M., Marmot, M., Goldberg, M., 2006. What does self rated health measure? Results from the British Whitehall II and French Gazel cohort studies. J. Epidemiol. Community Health 60 (4), 364–372. http://doi.

org/10.1136/jech.2005.039883.
Staudinger, U.M., Freund, A.M., Linden, M., Maas, I., 1999. Self, personality, and life regulation: facets of psychological resilience in old age. In: Baltes, P.B., Mayer, K.U. (Eds.), The Berlin Aging Study: Aging from 70 to 100. Cambridge University Press, New York, NY, US, pp. 302–328.
Stuck, S., Zuber, S., Kotte, M., Franzese, F., Gruber, S., Birkenbach, T., Operators, and all S. country team, 2018. Survey of Health, Ageing and Retirement in Europe. Release Guide 6.1.0. Munich Center for the Economics of Aging, Max-Planck Institute for Social Law and Social Policy, Munich.
Valensi, P., Lorgis, L., Cottin, Y., 2011. Prevalence, incidence, predictive factors and prognosis of silent myocardial infarction: a review of the literature. Archives of Cardiovascular Diseases 104 (3), 178–188. http://doi.org/10.1016/j.acvd.2010.11.013.
Vargas, C.M., Burt, V.L., Gillum, R.F., Pamuk, E.R., 1997. Validity of self-reported hypertension in the national health and nutrition examination survey III , 1988 – 1991. Prev. Med. 26, 678–685. http://doi.org/10.1006/pmed.1997.0190.
Vestergaard, S., Thinggaard, M., Jeune, B., Vaupel, J.W., McGue, M., Christensen, K., 2015. Physical and mental decline and yet rather happy? A study of Danes aged 45 and older. Aging Ment. Health 19 (5), 400–408. http://doi.org/10.1080/13607863.2014.944089.
Wang, Y., Freemantle, N., Nazareth, I., Hunt, K., 2014. Gender differences in survival and the use of primary care prior to diagnosis of three cancers: an analysis of routinely collected UK general practice data. PLoS One 9 (7). http://doi.org/10.1371/journal.pone.0101562.