

**Supporting Information for**

Going beyond the gender gap in healthy lifespans

Vanessa di Lego1\*, Marília R. Nepomuceno2, Cássio M. Turra 3

1 Wittgenstein Centre for Demography and Global Human Capital (IIASA, OeAW, Univ. Vienna), Vienna Institute of Demography at the Austrian Academy of Sciences

2 Max Planck Institute for Demographic Research, Rostock, Germany.

3 Universidade Federal de Minas Gerais, Cedeplar, Brazil.

\*Vanessa di Lego corresponding author

**Email:**  [Vanessa.DiLego@oeaw.ac.at](mailto:Vanessa.DiLego@oeaw.ac.at)

**This PDF file includes:**

Methods

Materials

Figures S1 to S9

Tables S1 to S3

SI References

Methods

1. Computing Disability- and Chronic-free life expectancy

We first estimate the disability-free life expectancy (*DFLE*) and the chronic-free life expectancy (*CFLE*) for ages 60 y and over using the Sullivan Method (1, 2).

The number of person-years lived free of disability () is calculated as,

Where  is the number of person-years lived without disability between ages and , is the total number of person-years lived in the age group and , and is the proportion of disabled individuals in the age group and . Then, life expectancy free of disability (*DFLE*) is calculated as:

The same is for chronic-free person-years lived, however with the prevalence for at least one chronic condition instead of prevalence of ADLs, and we call the person-years derived by the same process as . Hence, life expectancy free of chronic disease (*CFLE*) is:

The 95% confidence intervals for the *DFLE* and *CFLE* are computed directly from the standard errors derived from the weighted prevalence from each survey. For more details on how the weights are constructed for each survey, refer to data documentation on <https://g2aging.org/survey-overview>.

1. Gender Gap

The total gender gap in *DFLE* is calculated as:

And the gender gap in *CFLE* as:

In this way, larger gaps imply more advantage in favor of women and smaller gaps less advantage. Negative gaps imply that men have more advantage than women. Gaps are in absolute years.

1. Decomposing the gender gap

We later split the gender differences in and at ages 60 y and over into mortality and disability/chronic effects by five-year age groups. To decompose the gap, we apply the continuous change decomposition method, which assumes that covariates change continuously along an actual or hypothetical dimension, such as between two periods or between two populations (3–5). Such changes can be approximated by a linear combination of *n* partial derivatives of the function with respect to the covariates (3). Numerical integration is used to obtain the total contribution of the covariates for the variation of the aggregate measure. This method is very flexible, and can be used for decomposing gaps in different aggregate measures. For more details, refer to (5). Previous research has employed the methodology to estimate gaps in disability for Latin American and Caribbean (LAC) countries (6).

Materials

1. **Overall Health Data**

Health data for all countries is from the Gateway of Global Ageing, including doctor diagnosed chronic conditions and disability. We follow the recommendation by the Gateway of Global Ageing official report that for harmonization purposes, general diagnosis of chronic medical conditions (e.g. heart diseases), rather than a more specific condition (e.g. congestive heart failure), may be used to maximize the number of studies that may be included (7). We thus use the diagnosed chronic conditions surveyed for all HRS-sister studies and the following **disease variables:** RwDIABE, RwCANCRE, RwLUNGE, RwHEARTE, RwSTROKE, and RwARTHRE. There are indicator variables denoting whether or not the Respondent reports a doctor has ever told her/him that s/he had the specified condition. The conditions are 1) RwDIABE. diabetes or high blood sugar; 2) RwCANCRE. cancer or a malignant tumor of any kind except skin cancer; 3) RwLUNGE. chronic lung disease except asthma such as chronic bronchitis or emphysema; 4) RwHEARTE. heart problems, which include heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems; 5) RwSTROKE. stroke or transient ischemic attack (TIA); and 6) RwARTHRE. arthritis or rheumatism.

Specifically for the US, the Rand HRS 1992\_2018v1 files we use as basis for these doctor-diagnosed conditions, with the exception of cases that dispute a report from a prior wave, each of these variables is set to "yes" if the Respondent answered yes to the pertinent question in the current or any prior wave, and to "no" if the Respondent responded no at the current and all prior waves. We have constructed three new variables that capture prevalence of chronic conditions. “chronic” refers to having at least one of the conditions cited above. “chronic\_sum” refers to the total number of conditions ever diagnosed. “chronic\_severe” is a dummy variable that measures comorbidity, where an individual is diagnosed with having three or more of those conditions. **Robustness Checks.** At first, in order to evaluate further differences in onset of disease, we included the variable RADIAGDIAB, which indicates the age at which the respondent was first diagnosed with diabetes. RwRECCANCR indicates the most recent age at which the respondent was diagnosed with cancer. Respondents are asked the year in which they were most recently diagnosed with cancer, and these responses are converted to their age at diagnosis. Previous responses are carried forward if the respondent does not report a new cancer diagnosis. RwRECHRTATT indicates the most recent age at which the respondent had a heart attack. RAFRHRTATT indicates the age at which the respondent had their first heart attack. However, a first analysis showed that in the sample of HRS and other countries more than 75% was missing for age variables. So, we did not use these variables in the analysis.

**Country specific details. US (HRS).** We are using the Harmonized version B HRS: 37,495 observations. October 2018- There is a new updated version C, until 2019 that was updated now in 2022 and contains 42,233 observations. It is a Respondent level file so each row represents a unique Respondent. This leaves us with 18,747 observations using only wave 12 (year 2014) of HRS.

**Mexico (MHAS).** Version B.4 incorporates the latest released version of MHAS data, and adds several new variables. It contains 22,016 observations or rows- 22016. We are using the Harmonized VERSION B.4 (2001-2015), February 2022, for the MHAS data. The Mexican Health and Aging Study (MHAS) is a longitudinal household survey dataset for the study of health, economic position, and quality of life among the elderly. MHAS datasets as of September 2020. The MHAS (Mexican Health and Aging Study) Version B.4 incorporates the latest released version of MHAS data, and adds several new variables. It contains 22,016 observations or rows. It is a Respondent-level file so each row represents a unique Respondent. We will focus on Wave 4, which is for years 2014/2015. We will have 17,616 observations.

**England (ELSA).** We are using the Version G.2 (2002-2019), July 2021 for The English Longitudinal Study on Ageing (ELSA). It is a longitudinal household survey dataset for the study of health, economic position, and quality of life among the elderly (panel survey of people aged 50 and over and their partners, living in private households in England). Version G.2 incorporates the latest released version of ELSA data, which includes eleven main modules and the associated datasets, and adds variables and observations from Wave 9 with a total of 19,802 observations. It also adds new variables and makes adjustments and corrections. We will focus on Wave 7, nonetheless. The samples have been drawn from households which previously responded to the Health Survey for England (HSE). The seventh wave was conducted between June 2014 and May 2015 and included a refreshment sample selected from HSE 2011-2012.

**India (LASI).** The Longitudinal Aging Study in India (LASI) is a multidisciplinary, internationally harmonized panel study designed to be nationally representative of India’s population aged 45 and older. LASI is a joint project of three partnering institutions: International Institute for Population Sciences (IIPS), Harvard T.H. Chan School of Public Health (HSPH), and University of Southern California (USC). The first wave was conducted between 2017 and 2019 in 35 of India’s 36 states and union territories (except Sikkim). This initial sample, as released by USC, included 42,951 households and 72,262 individuals. The LASI sampling plan is complex and was based on the 2011 Indian Census with a multistage, stratified cluster sample design. The sample design includes three distinct selection stages in rural areas and four stages in urban areas. We use Version A.2 that makes corrections using the January 2021 released version of Wave 1 of the LASI data.

**Europe (SHARE).** This is Version F in the harmonized files and incorporates the latest released version of SHARE data, release 8.0.0, and adds observations from Wave 8. It contains 139,620 observations or rows. It is a Respondent-level file so each row represents a unique Respondent. It also adds new variables and makes adjustments and corrections. We focus on data from SHARE Wave 6, with the release 8.0.0 as of February 2022. SHARE uses a multistage stratified sample. Its weighting variables make its data representative of the target populations in constituent countries. Wave 6 does not still have full coverage of European countries, with the following countries only added in Wave 7: Finland, Lithuania, Latvia, Slovakia, Romania, Bulgaria, Malta and Cyprus.

**China (CHARLS).** The China Health and Retirement Longitudinal Study (CHARLS) is a longitudinal study of individuals over age 45 in China. Version D incorporates the latest released version of CHARLS data, and adds variables for Wave 4. It contains 25,586 observations or rows. It is a Respondent-level file so each row represents a unique Respondent; The sample population was selected as part of a stratified, multistage probability design. We will use Wave 3. As we concentrate on ages 50 and above due to the other samples we do not include individuals younger than 50. This leaves us with a sample size of 16,344 individuals.

**KLOSA- Korea.** The Korean Longitudinal Study of Ageing (KLoSA) is a panel survey of people aged 45 and over and their partners, living in private households in Korea. The survey elicits information about demographics, income, assets, health, cognition, family structure and connections, health care use and costs, housing, job status and history, expectations, and insurance. KLoSA surveys respondents every two years. Funded by the Korean Ministry of Labor, the Korean Institute of Labor (KLI) collected the first two waves, and the Korea Employment Information Service (KEIS) collected the Waves 3, 4, 5 and 6 of KLoSA, with the first wave of the KLoSA survey being conducted in fall/winter of 2006. The sample population was selected as part of a stratified, multi-stage area probability design. The first component of this sampling framework is the probability proportional to size (PPS) systematic sampling of the 2005 (South Korean) Census enumeration districts after stratifying by the location (15 major metropolitan cities and provinces) and characteristic of the district (urban or rural, and apartment building or non-apartment dwelling). Households were selected within PSUs from a listing of households in the Census identified as age-eligible; that is, inhabited by at least one person 45 years of age and older. This initial sample included 10,254 respondents age 45 and over. The second wave was conducted in 2008 and had 8,688 respondents. The third wave was conducted in 2010 and had 7,920 respondents. The fourth wave was conducted in 2012 and had 7,486 respondents. There was no refresher sample in Waves two through four. In 2014, a refreshment sample of individuals born in 1962 or 1963 was drawn and it included 920 individuals, which were added to the 7,029 remaining core sample respondents for a total of 7,949 Wave 5 respondents. The sixth wave was conducted in 2016 and had 7,490 respondents. We will focus on Wave 5. However, because we focus on ages >50, the sample is not 7,949. We use the hamornized Version C that contains 11,174 observations or rows. It is a Respondent-level file so each row represents a unique Respondent.

See Table S3 for the survey characteristics.

1. **Mortality data**

For mortality data, we use UN life tables from the 2022 Revision of World Population Prospects (United Nations 2022) for all countries with the exceptions of England, where the life tables are from the Office for National Statistics UK (ONS) estimates, as the ELSA study does not include Wales.

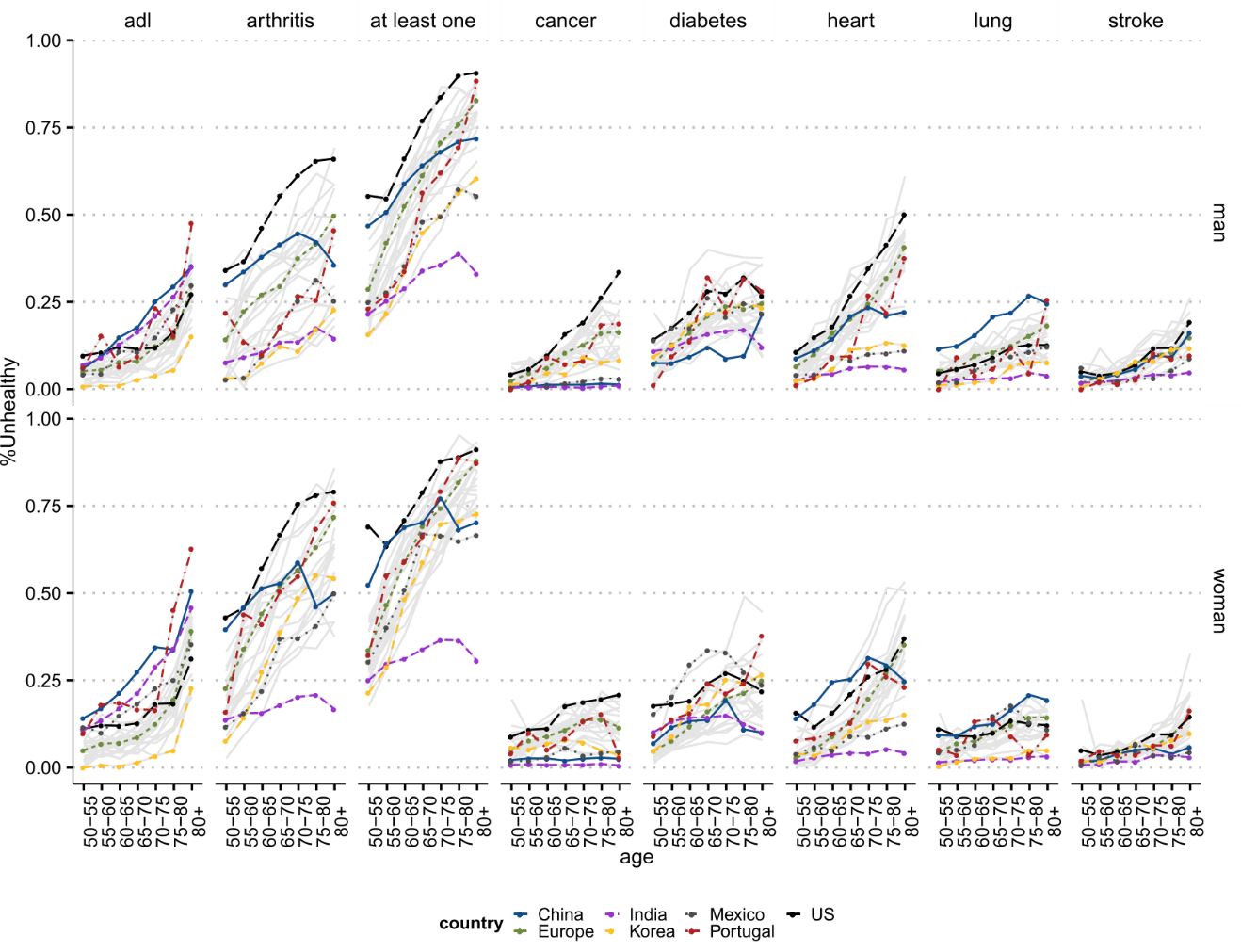


Fig. S1. Age-specific prevalence of health conditions for all countries (grey in the background) and selected countries for women and men. Notes. Panel “ADL” refers to the 5-item list of activities of daily living (ADLs), which include bathing, dressing, eating, getting in and out of bed, and using the toilet. Panel “At least one” refers to the constructed variable having at least one chronic doctor diagnosed diseases, which include diabetes, heart conditions, arthritis, cancer, stroke, and lung disease. Source: Gateway to Global Aging Data, Produced by the Program on Global Aging, Health & Policy, University of Southern California with funding from the National Institute on Aging (R01 AG030153).

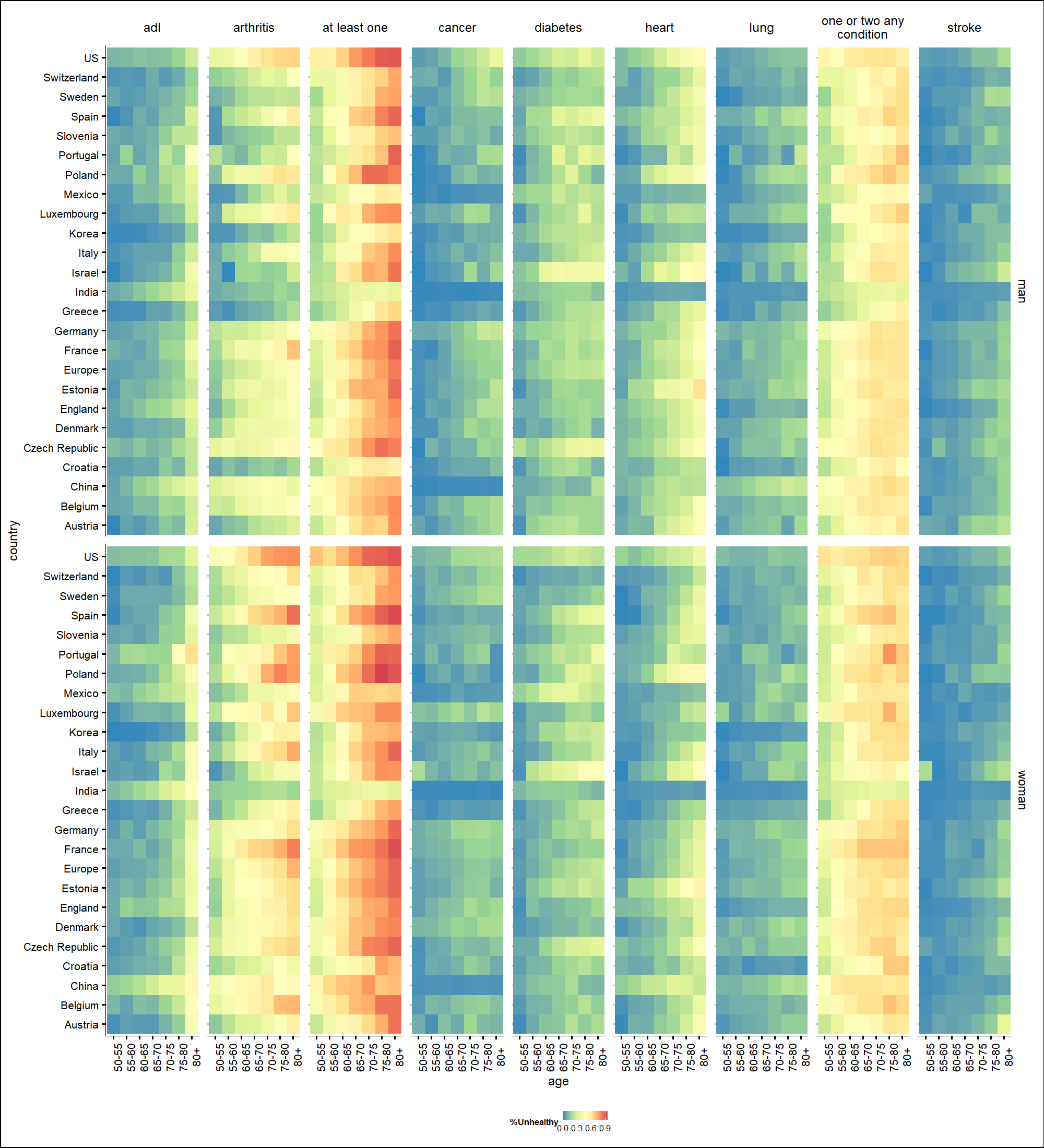


Fig. S2. Age-specific prevalence heatmap of health conditions for all countries for women and men. Notes. Panel “ADL” refers to the 5-item list of activities of daily living (ADLs), which include bathing, dressing, eating, getting in and out of bed, and using the toilet. Panel “At least one” refers to the constructed variable having at least one chronic doctor diagnosed diseases, which include diabetes, heart conditions, arthritis, cancer, stroke, and lung disease. Panel “One or two any condition” is similar, but aims at describing co-morbidities. Source: Gateway to Global Aging Data, Produced by the Program on Global Aging, Health & Policy, University of Southern California with funding from the National Institute on Aging (R01 AG030153).

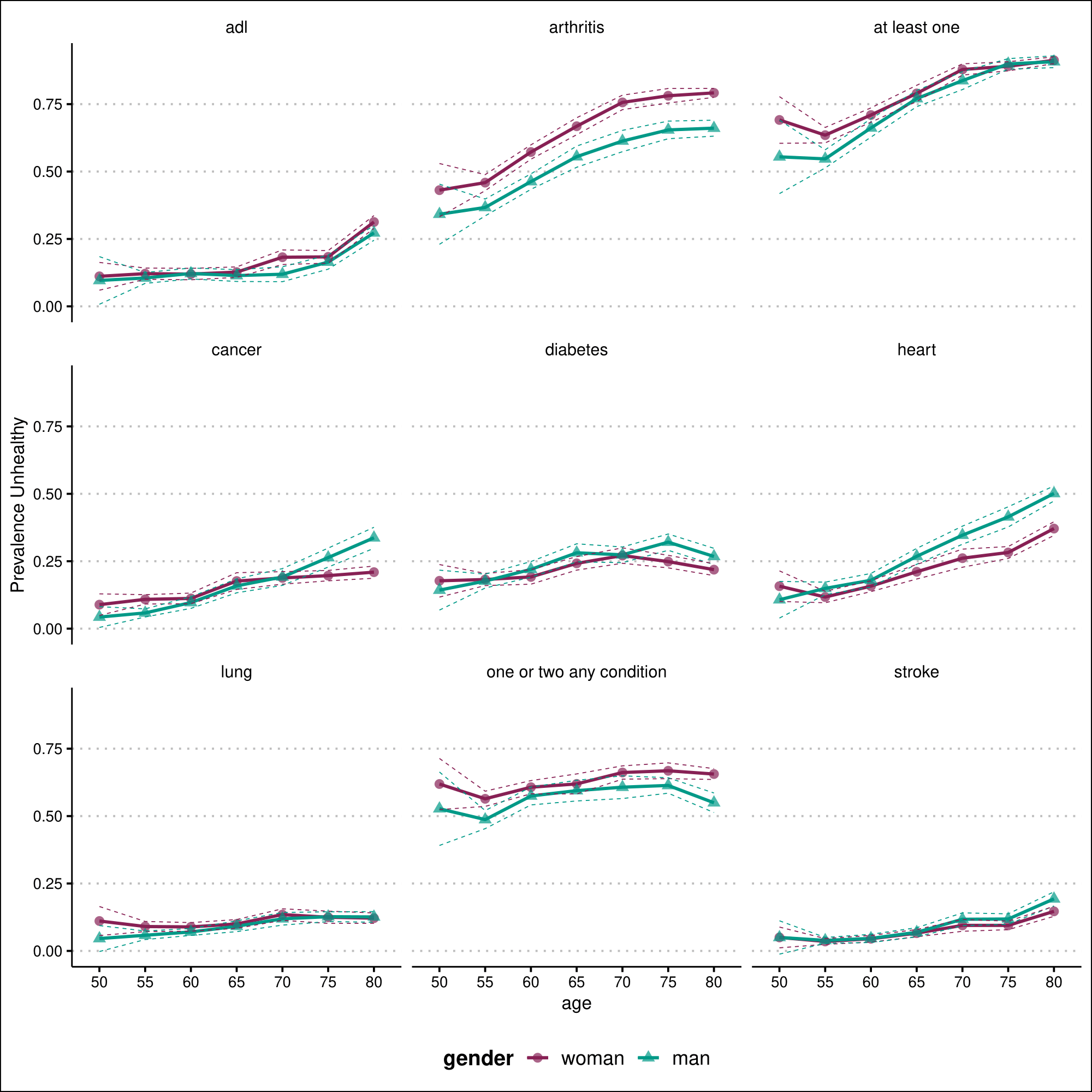


Fig. S3. Age-specific prevalence of health conditions for US (HRS), for women and men. Notes. Panel “ADL” refers to the 5-item list of activities of daily living (ADLs), which include bathing, dressing, eating, getting in and out of bed, and using the toilet. Panel “At least one” refers to the constructed variable having at least one chronic doctor diagnosed diseases, which include diabetes, heart conditions, arthritis, cancer, stroke, and lung disease. Panel “One or two any condition” is similar, but aims at describing co-morbidities. Source: Gateway to Global Aging Data, Produced by the Program on Global Aging, Health & Policy, University of Southern California with funding from the National Institute on Aging (R01 AG030153).

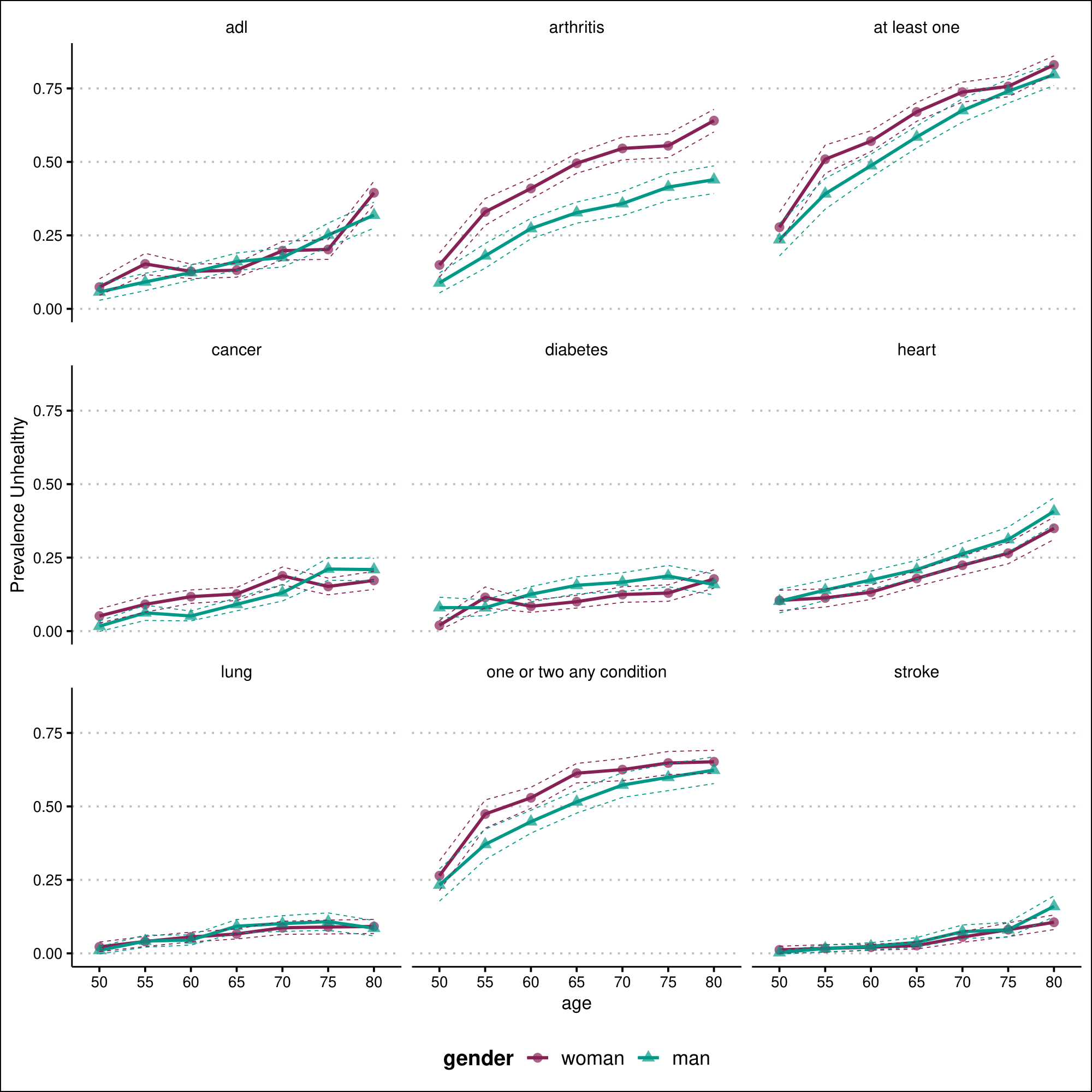


Fig. S4. Age-specific prevalence of health conditions for England (ELSA), for women and men. Notes. Panel “ADL” refers to the 5-item list of activities of daily living (ADLs), which include bathing, dressing, eating, getting in and out of bed, and using the toilet. Panel “At least one” refers to the constructed variable having at least one chronic doctor diagnosed diseases, which include diabetes, heart conditions, arthritis, cancer, stroke, and lung disease. Panel “One or two any condition” is similar, but aims at describing co-morbidities. Source: Gateway to Global Aging Data, Produced by the Program on Global Aging, Health & Policy, University of Southern California with funding from the National Institute on Aging (R01 AG030153).

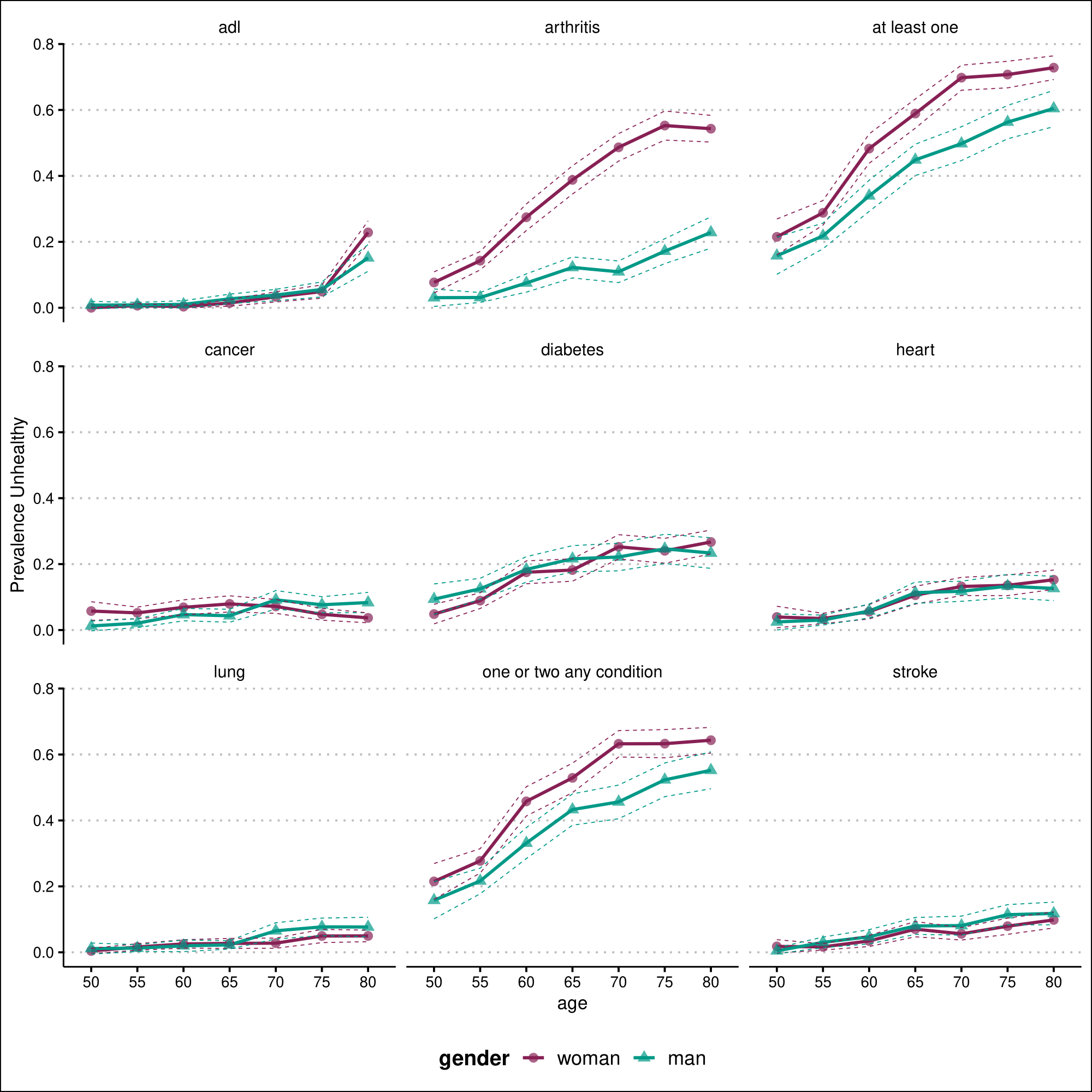


Fig. S5. Age-specific prevalence of health conditions for Korea (KLOSA), for women and men. Notes. Panel “ADL” refers to the 5-item list of activities of daily living (ADLs), which include bathing, dressing, eating, getting in and out of bed, and using the toilet. Panel “At least one” refers to the constructed variable having at least one chronic doctor diagnosed diseases, which include diabetes, heart conditions, arthritis, cancer, stroke, and lung disease. Panel “One or two any condition” is similar, but aims at describing co-morbidities. Source: Gateway to Global Aging Data, Produced by the Program on Global Aging, Health & Policy, University of Southern California with funding from the National Institute on Aging (R01 AG030153).

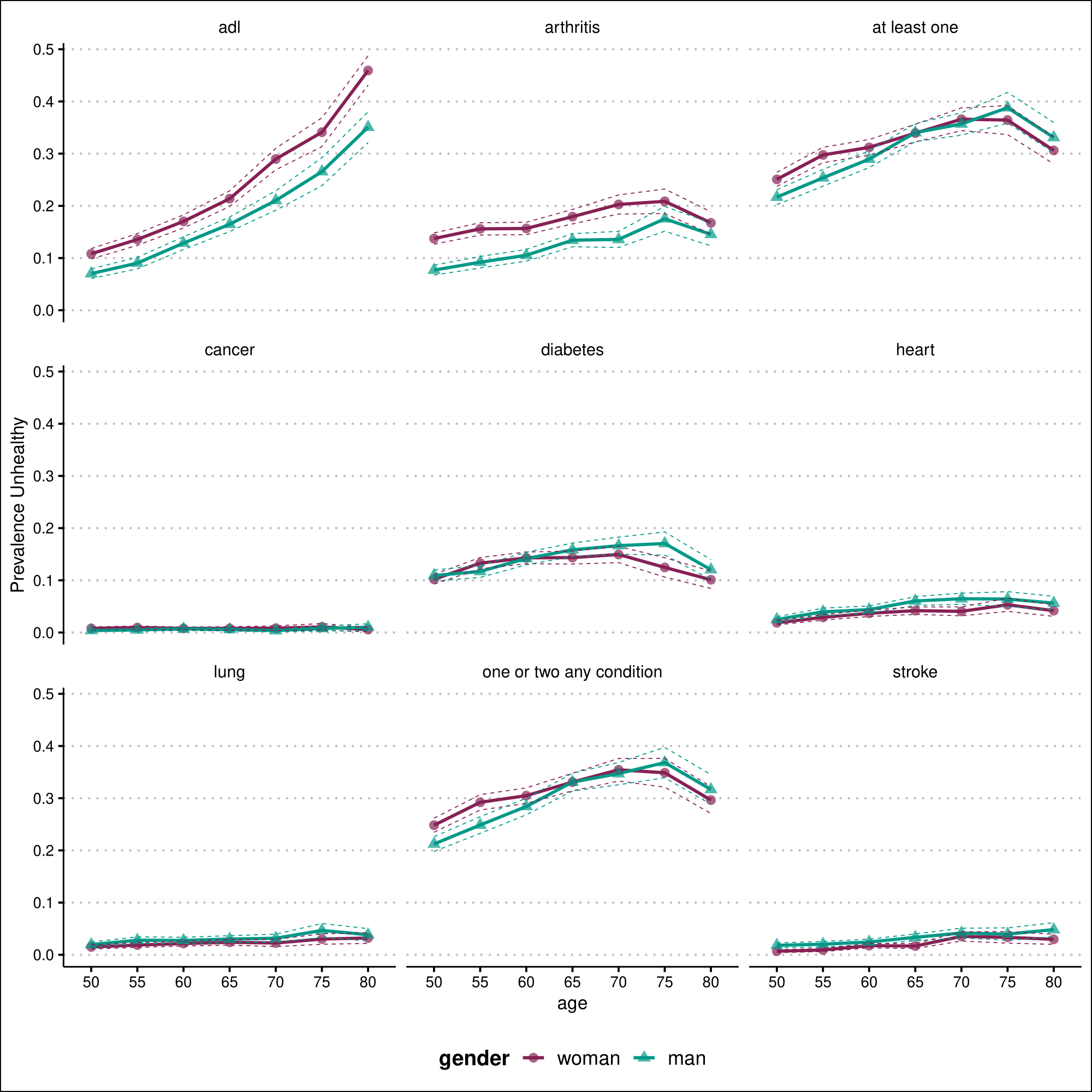


Fig. S6. Age-specific prevalence of health conditions for India (LASI), for women and men. Notes. Panel “ADL” refers to the 5-item list of activities of daily living (ADLs), which include bathing, dressing, eating, getting in and out of bed, and using the toilet. Panel “At least one” refers to the constructed variable having at least one chronic doctor diagnosed diseases, which include diabetes, heart conditions, arthritis, cancer, stroke, and lung disease. Panel “One or two any condition” is similar, but aims at describing co-morbidities. Source: Gateway to Global Aging Data, Produced by the Program on Global Aging, Health & Policy, University of Southern California with funding from the National Institute on Aging (R01 AG030153).

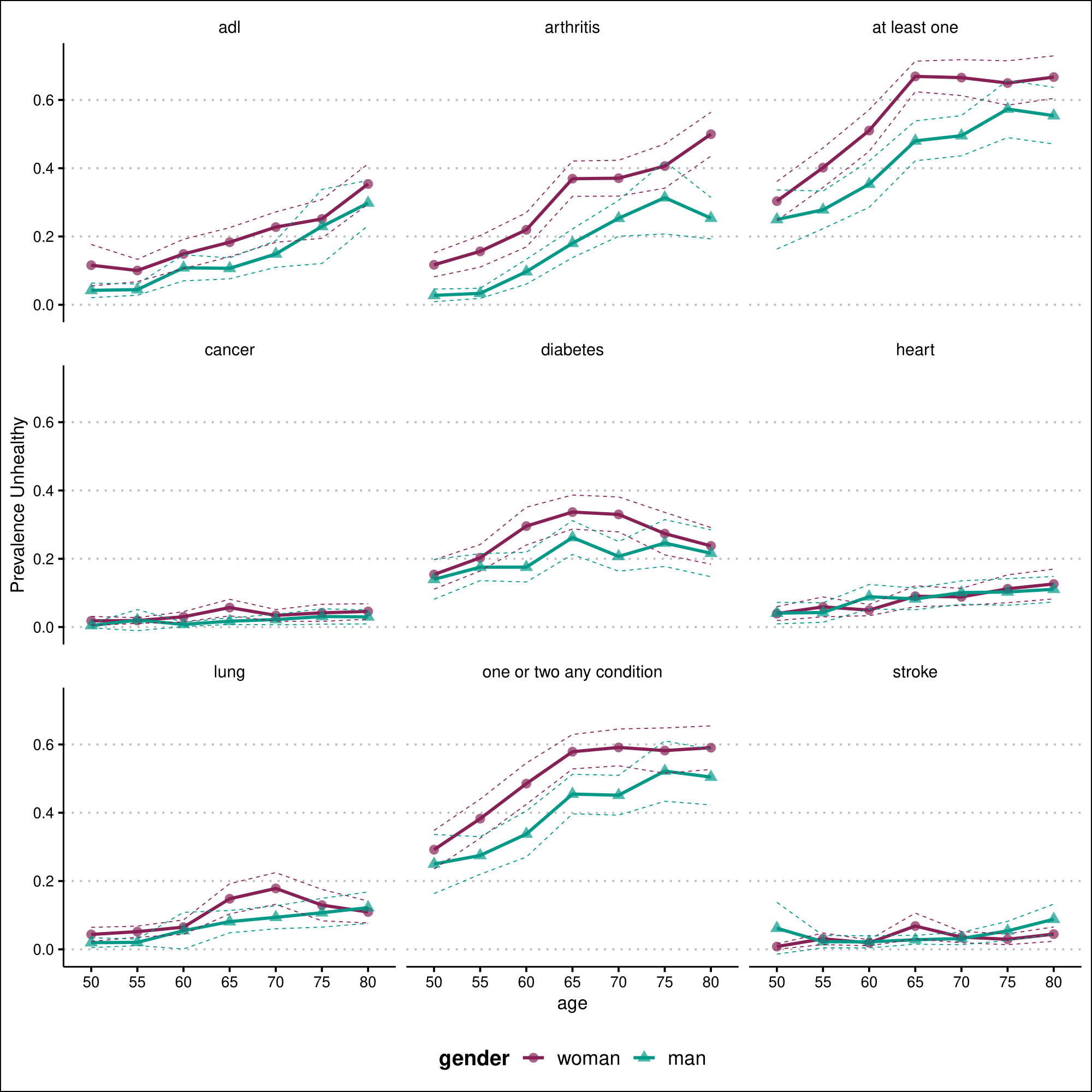


Fig. S7. Age-specific prevalence of health conditions for Mexico (MHAS), for women and men. Notes. Panel “ADL” refers to the 5-item list of activities of daily living (ADLs), which include bathing, dressing, eating, getting in and out of bed, and using the toilet. Panel “At least one” refers to the constructed variable having at least one chronic doctor diagnosed diseases, which include diabetes, heart conditions, arthritis, cancer, stroke, and lung disease. Panel “One or two any condition” is similar, but aims at describing co-morbidities. Source: Gateway to Global Aging Data, Produced by the Program on Global Aging, Health & Policy, University of Southern California with funding from the National Institute on Aging (R01 AG030153).

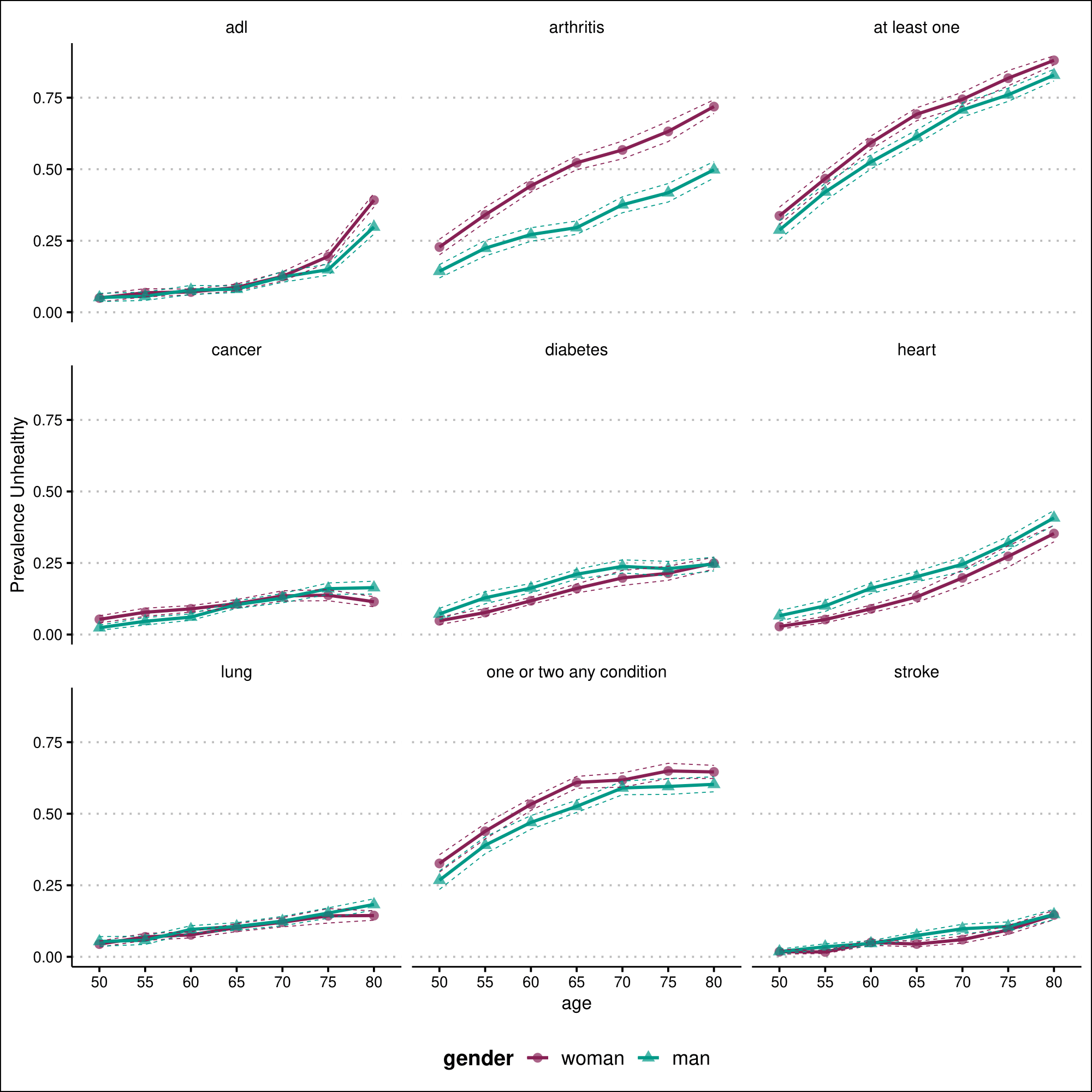


Fig. S8. Age-specific prevalence of health conditions for Europe (SHARE), for women and men. Notes. Panel “ADL” refers to the 5-item list of activities of daily living (ADLs), which include bathing, dressing, eating, getting in and out of bed, and using the toilet. Panel “At least one” refers to the constructed variable having at least one chronic doctor diagnosed diseases, which include diabetes, heart conditions, arthritis, cancer, stroke, and lung disease. Panel “One or two any condition” is similar, but aims at describing co-morbidities. Source: Gateway to Global Aging Data, Produced by the Program on Global Aging, Health & Policy, University of Southern California with funding from the National Institute on Aging (R01 AG030153).

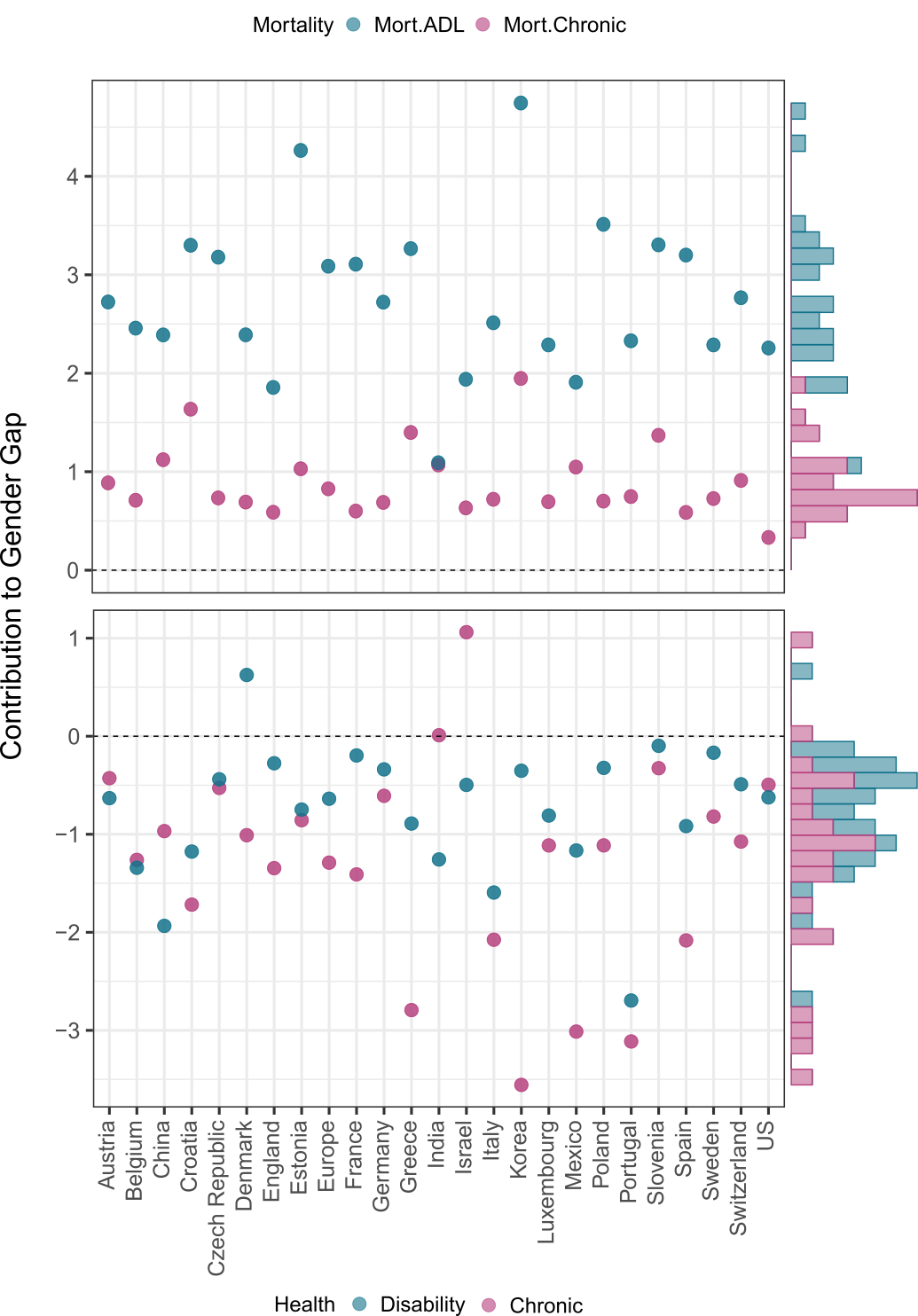


Fig. S9. Mortality and health components and their contribution to explaining the total gender gap. Source: Gateway to Global Aging Data, Produced by the Program on Global Aging, Health & Policy, University of Southern California with funding from the National Institute on Aging (R01 AG030153

Table S1. Decomposition of the gender gap (women-men) in disability-free life expectancy (DFLE) and life expectancy with disability (LEWD) at ages 60 y and over+ into mortality and disability effects by country, with 95% confidence intervals.



Source: Gateway to Global Aging Data, Produced by the Program on Global Aging, Health & Policy, University of Southern California with funding from the National Institute on Aging (R01 AG030153

Table S2. Decomposition of the gender gap (women-men) in chronic disease-free life expectancy (CFLE) and life expectancy with chronic disease (LEWD) at ages 60 y and over+ into mortality and chronic effects by country, with 95% confidence intervals.



Source: Gateway to Global Aging Data, Produced by the Program on Global Aging, Health & Policy, University of Southern California with funding from the National Institute on Aging (R01 AG030153

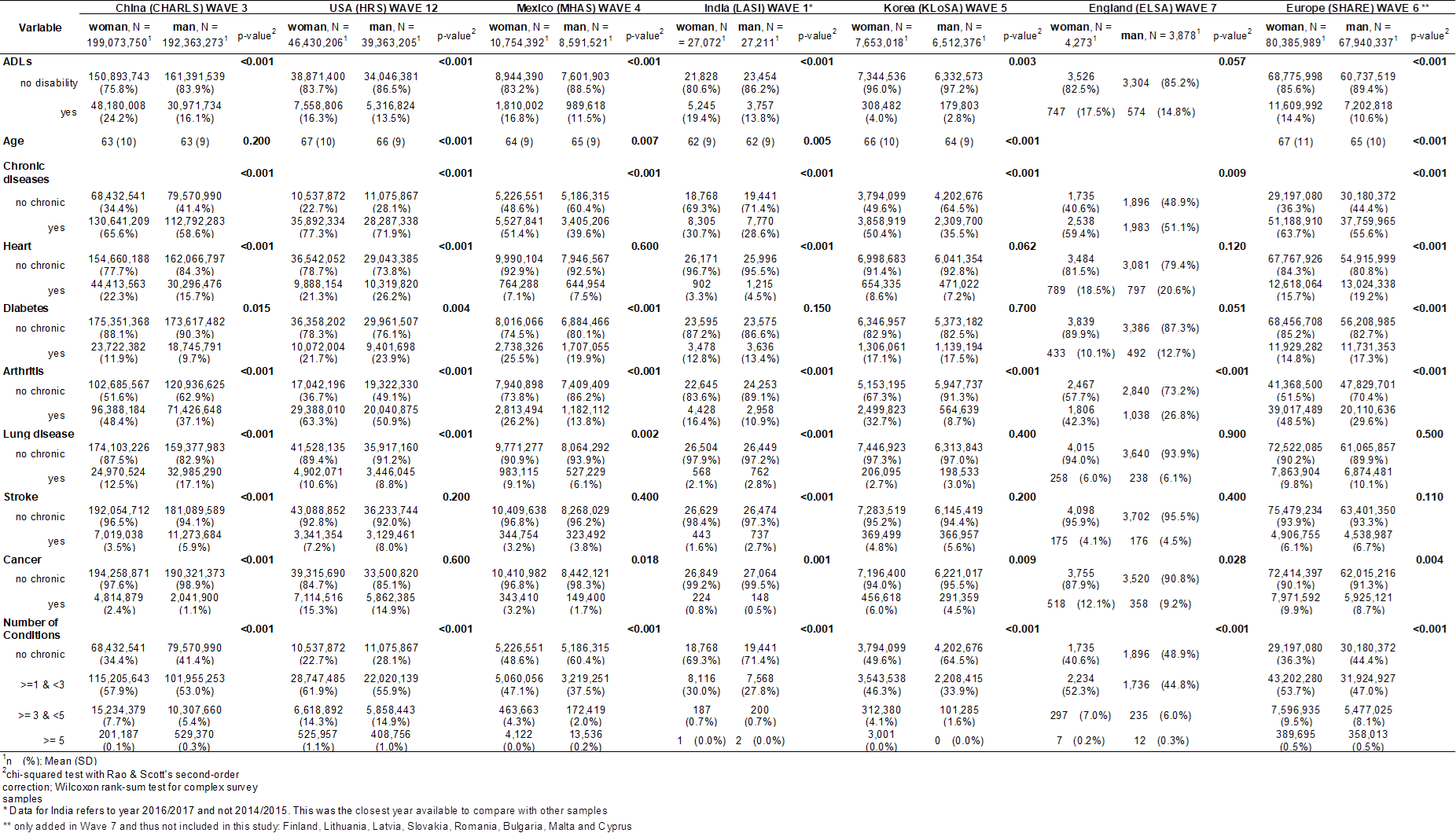


Table S3 Sample Characteristics and sample tests for different characteristics for women and men separately, weighted.

**SI References**

1. Y. Saito, J.-M. Robine, E. M. Crimmins, The methods and materials of health expectancy. *Stat. J. IAOS* **30**, 209–223 (2014).

2. D. F. Sullivan, A single index of mortality and morbidity. *HSMHA Health Rep.* **86**, 347–54 (1971).

3. S. Horiuchi, J. R. Wilmoth, S. D. Pletcher, A decomposition method based on a model of continuous change. *Demography* **45**, 785–801 (2008).

4. T. Riffe, Package “DemoDecomp” Type Package Title Decompose Demographic Functions (2018) https:/doi.org/10.1353/dem.0.0033 (October 15, 2019).

5. A. A. van Raalte, M. R. Nepomuceno, “Decomposing Gaps in Healthy Life Expectancy” in (2020), pp. 107–122.

6. M. R. Nepomuceno, V. di Lego, C. M. Turra, Gender disparities in health at older ages and their consequences for well-being in Latin America and the Caribbean. *Vienna Yearb. Popul. Res.* **19** (2021).

7. J. Lee, D. Phillips, J. Wilkens, Gateway to Global Aging Data: Resources for Cross-National Comparisons of Family, Social Environment, and Healthy Aging. *Journals Gerontol. Ser. B Psychol. Sci. Soc. Sci.* **76**, S5 (2021).