CESR-SCHAEFFER
WORKING PAPER SERIES

dornsife.usc.edu/cesr/   **healthpolicy.usc.edu**

Recommended Contents for Maximizing Harmonization Potential for the Health and Retirement Study and its International Network of Studies

Jinkook Lee, Jenny Wilkens, Drystan Phillips, David Knapp, Emma Nichols
University of Southern California

Preface

In preparing for population aging, the Health and Retirement Study was developed to provide an evidence base to study retirement transition and the health and wellbeing of older populations. Its scientific premises have attracted both scientists and policy makers, and under the leadership of the Division of Behavioral and Social Research at the National Institute on Aging, the HRS has served as the model for large-scale nationally representative longitudinal studies in a growing number of countries around the world.

In this paper, I aim to provide a summary of key contents in the HRS and its International Network of Studies (HRS-INS). In addition to key concepts and constructs, specific survey questions and response scales are also presented, providing guidance to maximize harmonization potential. For comparative research, comparable measures are pre-requisite, and the provision of these questions and response scales can help avoid unnecessary diversions, dampening the comparability.

Further, I will raise notable methodological considerations, which have stemmed from prior analysis or published work, as well as communication with data producers of existing surveys. These methodological issues can then be considered and addressed at the planning stage to the extent possible, avoiding common pitfalls.

This paper may also be useful for the principal investigators (PIs) of existing HRS-INS surveys. Whereas the PIs of existing surveys must be mindful of maintaining consistency to maximize their cross-wave concordance for longitudinal analysis, they may consider adding new constructs or survey questions to expand to new areas of research or to strengthen their existing protocol. Our recommended contents will provide guidance for such extensions.

Particular attention has been paid to strengthening the core cognition module for the HRS-INS that do not have an add-on Harmonized Cognitive Assessment Protocol (HCAP) study, although these recommendations can be also useful even for the HRS-INS with HCAP, as it will enable stronger linkage between the core HRS-INS and HCAP batteries, resulting in more robust cognition data for the entire sample.

Finally, while the HCAP was originally developed as an add-on study to the HRS-INS, with the success of the HCAP studies, there has been increasing interest in the development of new, stand-alone HCAP studies in the absence of ongoing longitudinal studies. For these new HCAP studies without a parent study, we suggest the core survey contents required for the investigation of key risk factors for dementia and the quantification of the costs of dementia.

The Health and Retirement Study (HRS) was originally established to replace the Social Security Administration's Retirement History Survey, but with a much greater emphasis on health and improved data collection on economic status (Suzman & Harper, 2013). Under the leadership of the National Institute on Aging, a new, multidisciplinary approach was taken, bringing academic researchers from diverse disciplines to collaboratively create and design the HRS to study the ways in which older adults' changing health interacts with social, economic, and psychological factors and retirement decisions (NIA, 2006). Over time, the objectives of the HRS have expanded to measure biological markers, additional psychosocial contents, and measures of cognitive functioning. The recent Harmonized Cognitive Assessment Protocol (HCAP) is a sub-study providing more detailed assessments of cognitive function and cognition related functional limitations, building on lessons from the Aging, Demographics, and Memory Study (ADAMS), which ran from 2000-2009. In this paper, I present the core contents with key concepts/constructs in the following order: demographics; health and health behavior; cognition; health and long-term care insurance and utilization; family, family transfers, and social networks, employment, retirement, and pension; wealth, income, and consumption; childhood experiences, stress, psychosocial measures, and end-of-life planning. The common questions asked across the HRS and its international network of surveys (HRS-INS) are summarized below, and exact question wording recommended based on cross-survey comparability are presented in the **Questionnaire for the International Network of Health and Retirement Studies**. Concordance tables showing the comparability of questions across the HRS-INS can be found on the Gateway to Global Aging Data website (https://g2aging.org/cas/home). Finally, I discuss interviewer observation that provides useful information about the interview and suggest a set of core survey questions for use in HCAP surveys without a parent study that are important for understanding the risk factors for and the burden of dementia.

1. **Demographics**

All HRS-INS collect basic demographic information about the survey respondents and often other household members.

1.1. **Birth Information and Gender.** All surveys collect the birth year and month of the respondents, and usually the age of other household members. Respondents are usually asked whether they were born in the country of interview, and if not, the actual country of birth, as well as when they came to the country of interview and their citizenship status for the country of interview. Gender is usually collected as male or female, though a few studies also include a transgender option.

1.2. **Education.** Educational attainment tends to be highly variable between countries and, as such, is collected in varying ways across the HRS-INS. Generally, the highest completed level of education is obtained, either categorically or through years of schooling. We recommend the use of categorical educational obtainment that can be mapped to International Standard Classification of Education (ISCED) categories (https://isced.uis.unesco.org/data-mapping/), as years of schooling cannot always be mapped to specific levels of education. We further recommend the use of the same educational categories for all individuals for which educational attainment is obtained in the survey, as well as maintaining consistency across waves.

1.3. **Marital Status.** Marital status commonly identifies those who have never been married, are married, partnered, separated, divorced, and widowed. Additionally, we recommend collecting the year the respondent was married or began living with their current partner, and if they have been married more than once, the year or age of their previous marriage(s), and how and when the marriage(s) was dissolved.

1.4. **Religion.** All studies collect the respondent's religious preferences. We have provided a generic listing of religions, but highly recommend tailoring the list to the local context. We further recommend asking about the importance of religion for those who declare having one.

2. Health and Health Behavior

All HRS-INS collect information about physical and mental health status and conditions by asking questions and conducting some objective performance tests, with an increasing number of studies collecting biological specimens. For biomarker collection protocol, we recommend consulting with the NIA-sponsored biomarker network's website: https://gero.usc.edu/cbph/network/protocol/.

2.1. **Physical and mental health**. Questions commonly asked to assess physical health include: (1) overall health status; (2) doctor-diagnosed health problems; (3) incontinence and hysterectomy; (4) sensory function; (5) dental health; (6) sleep; (7) falls and injuries; (8) pain and symptoms; and (9) depression.

2.1.1. *General health status*. First, self-rated general health status is a strong measure of overall health status and has been demonstrated to correlate with subsequent health service use, functional status, and mortality (CDC, 2000). All HRS-INS included a self-rated general health question, but using two different scales based on cultural consideration of response patterns, either ranging from poor to excellent or ranging from very poor to very good. Second, several studies (HRS, ELSA, SHARE, TILDA, LASI, CHARLS, and KLoSA) also ask a follow-up question on whether any health problem or disability limits the kind or amount of paid work one could do. A more limited set of surveys (SHARE, TILDA, and JSTAR) ask more generally about whether a health problem or disability limits activities in any way. Methodological issue: Potential subjective bias associated with self-reports has been recognized, and, to overcome such bias, several HRS-INS (e.g., HRS, ELSA, SHARE, LASI) included anchoring vignette questions that describe the health status of hypothetical persons and ask respondents to evaluate the health of those persons using the same scale that they used to describe their own health. Vignettes have been applied successfully on international comparisons of health and work disability (Kapteyn, Smith, and Van Soest, 2007).

2.1.2. *Doctor-diagnosed health problems*. All HRS-INS ask questions on doctor-diagnosed chronic diseases and other health problems. Respondents reporting such a diagnosis are asked when they were first diagnosed, whether they are receiving medical treatment, and in follow-up interviews for some studies, whether the condition has worsened. The chronic conditions most commonly asked about include: hypertension, diabetes, cancer, chronic lung disease, asthma, heart disease, stroke, hypercholesterolemia, arthritis or rheumatism, psychiatric problems, and memory-related disease such as Alzheimer's disease or dementia. A more limited number of surveys included questions on the following additional conditions that are common in their local settings: stomach or other digestive diseases, osteoporosis, Parkinson's disease, liver disease, and kidney disease. Methodological issue: It is important to note that the type of healthcare professionals allowed to have made a diagnosis varies from study to study. HRS and other high-income countries specifically exclude diagnoses made by nurses/nurse practitioners and chiropractors, but low-and-middle income countries like CHARLS and LASI allow diagnosis by nurses, practitioners of traditional medicine, and other healthcare professionals. There is additional variability in the wording of chronic lung disease, with some studies specifically excluding asthma and other studies explicitly including asthma.

2.1.3. *Incontinence and hysterectomy*. Most HRS-INS asked whether the respondent experienced urinary incontinence, and female respondents were also asked about menstruation or menopause, as well as having a hysterectomy.

2.1.4. *Sensory function*. Most HRS-INS asked respondents to rate their eyesight in general, and only a handful of surveys asked for two separate ratings of distance and near vision, either in addition to or in place of general vision. Respondents were also frequently asked whether they usually wear glasses or contact lenses and whether they have had cataract surgery. LASI and SAGE conducted a visual acuity test, using the Tumbling E logMAR chart on a laptop for a distance of 3 meter and 40 cm from the screen to measure distance and near vision (Ehrlich

et al., 2022). Most HRS-INS also asked for a rating of overall hearing ability and whether a hearing aid is usually used. Similar to self-rated health, two different scales are used for self-rated vision and hearing. For those who reported poor vision and hearing, we suggest asking about when vision and hearing loss occurred.

2.1.5. **Dental health**. Most HRS-INS asked whether the respondent has lost some or all of their natural permanent teeth and whether they usually wear dentures, including implants.

2.1.6. **Sleep**. Common questions asked about sleep include: how often the respondent had trouble with (1) falling asleep, (2) waking up during the night, (3) waking up too early and not being able to fall asleep again, and (4) whether they are taking any medication to help them sleep. It can be helpful to differentiate between over-the-counter and prescription sleep medications.

2.1.7. **Falls and injuries**. Falls and other injuries are prevalent among older adults and, therefore, asked in all HRS-INS, most commonly for the time window of the past two years. If the respondent had a fall or injuries, follow-up questions are asked on how many times a fall was experienced, whether the injury was serious enough to need medical treatment, as well as whether it involved hip fracture. Several surveys in East Asia, like CHARLS, KLoSA, and LASI, also asked about traffic accidents and natural disasters. Although not commonly asked, we recommend including a question on brain injury, as it has been identified as an important risk factor of dementia (Livingston et al., 2020).

2.1.8. **Pain and symptoms**. For those who reported pain, the severity of pain, whether pain limits any activities, and whether pain medication is taken to control the pain were asked in most surveys. In measuring the severity of pain, a 3-point Likert scale, ranging from mild to severe, was commonly used. It will be beneficial to differentiate between over-the-counter and prescription pain medication given the increasing reliance on opioids. Several HRS-INS (namely HRS, MHAS, SHARE, and LASI) included additional questions on persistent symptoms, the most common being: (1) persistent swelling in legs, ankles or feet; (2) shortness of breath while awake; (3) persistent dizziness or lightheadedness; (4) back pain; (5) persistent headaches; (6) severe fatigue or exhaustion; (7) persistent wheezing, cough, or bringing up phlegm.

2.1.9. **Depressive symptoms**. Most HRS-INS asked about depressive symptoms, drawing questions from the CES-D (Radloff, 1977). Methodological issue: Different scales have been employed in different surveys, making cross-survey comparisons difficult. HRS, MHAS, and ELSA used a binary scale of CES-D, whereas KLoSA, JSTAR, TILDA, CHARLS, and LASI used a 4-point Likert scale. Additionally, the number of items frequently differ between surveys or waves. SHARE used the Euro-D. Given the lack of consistency, validated cut-offs for the presence of depression would be extremely helpful for cross-survey analysis.

2.2. **Functional health**. All HRS-INS included questions on limitations (that last 3+ months) with mobility, activities of daily living (ADL), and instrumental activities of daily living (IADL).

2.2.1. **Mobility**. The most commonly included questions ask about difficulty with: (1) walking a short distance; (2) sitting for about 2 hours; (3) getting up for a chair after sitting for long periods; (4) climbing several flights of stairs without resting; (5) climbing one flight of stairs without resting; (6) stooping, kneeling, or crouching; (7) reaching or extending arms above shoulder level; (8) pulling or pushing large objects like a living room chair; (9) lifting or carrying objects that weigh over 5 kg, like a heavy bag of groceries; and (10) picking up a small coin from a table. JSTAR and MARS ask about difficulty climbing several steps or one step without a handrail, while most other surveys ask about climbing several flights or one flight of stairs.

2.2.2. **ADL**. The most common ADL questions ask: "Because of a health or memory problem do you have any difficulty with …": (1) dressing, including putting on shoes and socks; (2) walking across a room; (3) bathing or showering; (4) eating, such as chewing and cutting up food; (5) getting in or out of bed; and (6) using the toilet, including getting up or down (or

squatting). <u>Methodological issue</u>: While the ADL questions ask about 'having difficulty,' many researchers have attempted to use the responses to these questions to assess 'care needs.' However, respondents who have difficulty with ADLs may not need help, and in such cases, there would be no care need. Considering such scenarios, response options for this set of questions have been adjusted to separate out those having difficulty but not needing help from those having difficulty and needing help.

2.2.3. *IADL*. Similar to ADLs, IADL questions ask about having difficulty most commonly with the following activities: (1) preparing a hot meal; (2) shopping for groceries; (3) making telephone calls; (4) taking medications; (5) doing work around the house or garden; and (6) managing money, such as paying bills, keeping track of expenses. <u>Methodological issue</u>: In addition to the methodological issue raised above regarding having difficulty and needing help, it is important to acknowledge that there might be some subpopulations who have never engaged in such activities for cultural reasons. For example, in many Asian cultures, men are not expected to prepare a hot meal.

2.2.4. *Care receipt*. All HRS-INS ask about help received for ADL and IADL difficulties, mostly following up for those who reported any difficulties. As noted, the distinction between having difficulties and needing help may be required. Furthermore, there are differences across studies and waves for the grouping of activities respondents receive help for and the number of caregivers the respondent can report. We recommend asking help received for ADL separately from IADL, as well as questions about the care providers, including the relationship with the primary care providers, frequency and duration of care received, and whether their care needs are met. We further recommend placing these questions in the Functional Health module following the ADL and IADL questions.

2.3. Physical assessment

Methodological issues arise, as there are some subtle and some significant differences in the exact protocol followed or equipment used during the physical assessment (Wang, Wilkens, and Phillips, 2023).

2.3.1. *Anthropometric measurements.* Most HRS-INS measure weight and height, and several studies also asked respondents to self-report height and weight. Other common anthropometric measurements are waist and hip circumference.

2.3.2. *Blood pressure and pulse.* The most common protocol followed by HRS-INS was to measure systolic and diastolic blood pressure and pulse 3 times at a sitting position with about 1-minute intervals between the measurements.

2.3.3. *Walking speed.* Several HRS-INS administered a timed walk test, asking respondents to walk a short distance twice on a flat surface at a normal speed. The distance walked has been variable: about 2.5 meters was used in four studies, 3 meters was used in two studies, and 4 meters was used in one study. Additionally, this test was often administered to a subsample, either random (HRS Wave 7), aged 60+ (ELSA, CHARLS), aged 65+ (HRS Wave 8) or aged 75+ (SHARE).

2.3.4. *Grip strength*. Most HRS-INS measured grip strength, using a hand-held dynamometer. The most common protocol asked respondents to hold the dynamometer at a 90-degree angle and squeeze the handle for several seconds in a standing-up position, and two measurements were taken with each hand.

2.3.5. *Balance test*. Among all HRS-INS, only HRS, ELSA, CHARLS, and LASI measured balance in three positions: semi-tandem, full-tandem, and side-by-side. <u>Methodological issue</u>: The sequence of positions for the balance test differs across surveys, and the subsequent test is conditional on the respondent's success with the initial stance(s).

2.4. **Health behavior**. The HRS-INS are increasingly collecting information about preventative care, and all HRS-INS collected information on smoking, drinking, and physical activities.

2.4.1. ***Preventative care***. Several HRS-INS asked about various screening tests, including: (1) flu shots, (2) cholesterol tests, (3) pneumonia vaccines, (4) colon cancer screenings, (5) mammograms, (6) pap smears, and (7) prostate exams. Again, the time reference period varies significantly across studies.

2.4.2. ***Smoking.*** Most surveys asked about the respondent's smoking history: whether they ever smoked, the age or year when they started smoking, and the age or year when they stopped smoking. For current smokers, the quantity of smoking is asked. Most surveys ask only about cigarettes, but limited studies ask about other tobacco products or the use of smokeless tobacco.

2.4.3. ***Drinking.*** Most surveys asked whether the respondent has drunk alcohol within a specified time period, and the frequency and amount of alcohol intake for drinkers. Several surveys (HRS, SHARE, MHAS, LASI, ELSI, and MARS) also asked about binge drinking and 4-item questions drawn from the CAGE questionnaire for the detection of alcoholism (O'Brien, 2008). Few surveys ask whether the respondent has ever drunk alcohol.

2.4.4. ***Physical activity***. Most surveys asked about the frequency of different levels of physical activity, including vigorous, moderate, or mild physical activities. <u>Methodological issue</u>: The examples of activities given for each level of physical activity are variable across studies, as are the scales or manner of collection for the frequency of participation. Furthermore, limitations of self-reported physical activity have been widely recognized (Kapteyn et al., 2018), and an increasing number of studies have embraced the direct measurement of physical activities and sleep using mobile sensors.

3. **Cognition.** All HRS-INS included questions to measure cognitive function, and when respondents are incapable or refuse to participate in cognitive tests, several surveys (HRS, ELSA, SHARE, MHAS, LASI, ELSI) administered a proxy interview by asking a family member or friend about the respondent's cognitive status.

Prior work has shown that the cognitive assessments in the HRS-INS represent comparable underlying cognitive constructs (Kobayashi et al., 2021). Our team has conducted extensive analysis of core cognitive tests as well as newly available data from the Harmonized Cognitive Assessment Protocol (HCAP) (Gross et al., 2023; Nichols et al., 2022, Nichols, Jones et al., 2024), and our recommendations for the cognition module are based on weighing a combination of factors, including consistency over time with prior waves, consistency across countries, appropriateness across settings, ability to provide linkage with the HCAP battery, and utility of items in contributing to measurement precision across the range of cognitive functioning (i.e., if the cognitive test items provide information about those with both severe and mild impairment). Although some of these items have been included in prior HRS-INS batteries, we recommend not including items with low variability (i.e., object or person naming) and scoring difficulty (e.g., backward count), and instead taking advantage of better performing cognitive tests from the core (self-reported memory rating, word recall, time orientation, animal naming, serial 7s) and HCAP (i.e., 3-stage task, symbol cancellation, and backward day naming). Given the focus on strengthening the core cognition module for HRS-INS without an add-on HCAP study, we also include additional tests that we recommend for the creation of an *extended* version of the cognitive battery.

<u>Methodological issues</u>: It would be helpful to include interviewer observations from the administration of the cognitive tests, asking interviewers to report any factors that may have impaired the respondent's performance on the test, such as poor vision, poor hearing, extreme fatigue, physical impairment that affects their ability to perform the test, interruption or distraction, noisy environment, problem with the laptop, language barrier, and being under the influence of alcohol. In coding missing data, responses of "don't know" should be clearly

differentiated from "refusals" or general missing data codes. This allows a response of "don't know" on a cognitive test to be re-coded as a score of 0.

Some decisions around which tests to include may be impacted by the planned mode of administration; for example, though we recommend the use of symbol digit cancellation, it is not feasible to administer this test via telephone. Given the shifts towards web-based rather than telephone-based assessments, we decided not to limit our recommendations to items that could be administered over the telephone. With any differences in mode, randomization or calibration studies are needed to quantify biases due to mode differences (Smith et al., 2023; Domingue et al. 2023). The recommended cognition items are as follows:

3.1. ***Self-reported memory rating*** and changes in memory since the last interview.
3.2. ***Sentence reading or speaking.*** This task asks the respondent to write a simple sentence, or to say a simple sentence if they are unable to write one. This is a simple assessment of language and praxis.
3.3. ***Word recall***. Word recall is a strong measure of memory performance. Most HRS-INS read a set of 10 words once for the core interview and 3 times in the HCAP interview. The recommended time interval between immediate and delayed recall tests. For the HCAP interview, the recommended time interval for word recall tests is about 20 to 25 minutes. We recommend using only auditory cues, with amplification to accommodate hearing impairment, to improve comparability to LMIC settings where visual cues are infeasible due to high levels of illiteracy. <u>Methodological issue</u>: In selecting specific words for this test, familiarity with the word in the particular language and culture needs to be considered.
3.4. ***Time orientation***. Day of the month, month of the year, year, and day of the week are most commonly asked. Though most respondents will answer these items correctly in population-based studies, incorrect answers signal disorientation and potentially more severe cognitive impairment.
3.5. ***Overlapping pentagons***. Copying figures is a measure of visuospatial functioning. Respondents are asked to copy a drawing of two overlapping pentagons. Up to two attempts are allowed in administering this test. <u>Methodological issue:</u> Scoring needs care, as some studies require angles to be maintained in addition to requiring that two five-sided shapes intersect to form a four-sided shape, making the task comparatively more difficult. Additionally, there can be potential interrater differences in scoring which should be addressed during interviewer training and/or survey administration or scoring.
3.6. ***3-stage task***. This test involves following an instruction to receive a paper, fold it, and either return it or place it somewhere. This test involves language capacity and executive functioning but is typically considered as a language assessment. This is incorporated in some core studies and is also administered in HCAP.
3.7. ***Symbol cancellation***. Symbol cancellation is a strong measure of executive functioning. This is also a common test administered in HCAP in LMICs. Though the letter cancellation version is more often used in high-income settings, symbol cancellation is feasible and acceptable across all settings, making it an excellent candidate for cross-national comparisons. Though sparingly used in core HRS-INS to date, adding symbol cancellation would bolster assessments of executive functioning, improve links between the core HRS-INS and HCAP, and facilitate cross-national comparisons. <u>Methodological issue:</u> To maximize comparability, care is needed to make sure the exact same form, with the same symbol size and spacing is used between studies.
3.8. ***Animal naming***. Animal naming is a test of language fluency. Respondents are asked to name as many animals as they can within 60 seconds. This task is commonly used both in core HRS-INS and HCAP studies.

3.9. ***Serial 7s***. Serial 7s is a test of executive functioning. Respondents are asked to count backwards from 100 by 7. This task is commonly used in the core cognition module and in HCAP. Methodological issue: In LMIC, innumeracy can be an issue. In these settings, a missing data code for "cannot count" should be added, and imputation should be considered for these responses.

3.10. ***Backward day naming***. Backward day naming is a test of executive functioning. Respondents are asked to name the days of the week backward, starting from Sunday. Backward day naming is recommended as a replacement for backward counting (which is used in a number of HRS-INS) because it is more easily administered in places where the innumeracy rate is high.

3.11. ***Story recall [extended version]***. The Brave Man story recall task is an assessment of logical memory that asks respondents to recall 6 different points of a story immediately after hearing it, and after a delay. Though this test has not previously been used in the HRS-INS, it has performed well in HCAP studies. Methodological issue: some story points may need to be adjusted to ensure cultural relevance in novel settings.

3.12. ***Constructional praxis [extended version].*** The constructional praxis test evaluates the respondent's ability to see a shape and copy it. The test includes 4 shapes. There is also a delayed test that asks respondents to draw the earlier shapes to the best of their ability. The initial assessment is a visuospatial functioning test, and the delayed recall portion is a test of memory. Methodological issue: Similar to the assessment of overlapping pentagons, scoring needs care, and assessment of inter-rater differences is useful.

3.13. ***Go-No-Go [extended version].*** The Go-No-Go test is a test of executive functioning with two parts. In the first part, respondents are asked to repeat a specific tapping pattern in response to the interviewers taps, in the second part, the exercise is repeated but with a rule change. Methodological issue: Motor issues can pose problems with test administration.

3.14. ***Proxy Cognition***. Informants were first asked to provide an overall rating of respondent's memory and cognitive capacity and its changes in the past two years and then asked to answer questions from IQCODE (Jorm & Jacob 1989) about the changes in cognition in the past two or ten years. While two years is more commonly used in existing surveys, ten years may be more desirable given the varying periodicity across survey waves and its use in the original IQCODE questionnaire. Additional questions are also included, asking whether respondents ever get lost in a familiar environment, wander off and do not return, can be left alone for an hour, or have visual or auditory hallucinations. Methodological issue: Given evidence that informant characteristics can impact reporting (Nichols, Gross et al. 2024), these data should be captured and considered in the interpretation of reporting from informants.

4. Health and Long-Term Care Insurance and Utilization

All HRS-INS surveys asked about public and private health insurance, as well as healthcare utilization and expenditures, including hospitalization, doctor visits, medication, dental care, and long-term care. A handful of surveys also asked about long-term care insurance and nursing home stays.

4.1. **Insurance policies.** Studies differ in what is asked regarding insurance coverage because the structure of insurance varies substantially based on a country's public health insurance system and social security policies. Studies of countries with universal national health insurance typically do not need to ask about public health insurance coverage or premiums as this information is shared across the population. However, the existence of national health insurance does not typically eliminate a study's need to ask about insurance policies. In some countries with national insurance, the public insurance only covers a percentage of costs (e.g., in France, the main public health insurance covers 80% of costs) and it is common for people to purchase supplementary (also known as complementary) insurances that augment coverage or pay for costs not fully covered by the national plan. In countries where health insurance coverage is

voluntary, questions about insurance coverage and cost are essential to understanding individual choices. In these cases, insurance policies are more varied, and so survey's need additional questions to discern a private policy's coverage and care use costs.

Given the varied nature of country insurance systems, we encourage studies to collect basic information on all of the respondent's insurance policies, including (1) covered services, (2) provider (e.g., public or private), and (3) premium payment and payment frequency. At a minimum, studies should ask whether a respondent has access to health insurance that provides coverage for costs associated with hospitalization, outpatient care services like primary and specialist care, and medications. Depending on the study's interests, the scope of the insurance policies module can be expanded to health-adjacent topics (e.g., long-term care insurance, sickness, accident, or disability insurance, life insurance). Most studies ask questions about life insurance, including the beneficiary and death benefit amount. In some countries, help with care-related costs may be provided through other types of insurance, such as sickness or disability insurances.

The basic information asked about any type of insurance is often expanded to include questions about how the individual has access to an insurance policy (e.g., through an employer or spouse) or whether the insurance coverage is required by the government, required or provided by an employer, or entirely voluntary. In countries where it is common to go without health insurance coverage, studies will also ask why a person does not have health insurance coverage.

4.2. **Healthcare utilization**. Most of the HRS-INS asked about the utilization of various types of healthcare and associated out-of-pocket expenditures. <u>Methodological issue</u>: The frequency of healthcare utilization and the amount of associated out-of-pocket expenditure is asked based on various reference time periods, ranging from the past year, past 2 years, and last 4 weeks, to the last visit. This difference in reference time periods makes cross-survey comparison difficult. We recommend the reference period of the past year, as it is most common across existing HRS-INS surveys, and most burden of disease studies report annual spending. For more accurate estimates, linkage to healthcare records is recommended.

4.2.1. *Hospitalization.* Common questions on hospitalization include: (1) whether the respondent was ever hospitalized during the reference period; and for those who reported hospitalization, (2) the number of hospitalizations, (3) the number of nights they stayed at the hospital, and (4) out-of-pocket costs for hospitalization.

4.2.2. *Outpatient care*. Most HRS-INS asked about the frequency of use and the associated out-of-pocket costs separately for (1) visits to the doctor's office and (2) other outpatient care or surgery separately. In addition, several surveys in Asia included additional questions on traditional healthcare.

4.2.3. *Medication*. Several HRS-INS (including HRS, SHARE, ELSI, KLoSA, CRELES, SPS) asked whether respondents regularly take any prescription medication and the out-of-pocket costs for the medication.

4.2.4. *Dental care*. Most HRS-INS asked whether respondents received dental care during a certain reference period and the out-of-pocket cost of dental care.

4.3. **Long-term care utilization.** Long-term care (LTC) covers home care, assisted living, adult daycare, respite care, hospice care, or stays in nursing homes or residential care facilities. LTC insurance helps cover the cost of long-term care, but often does not cover it completely. Reflecting country differences in policies for providing LTC, surveys vary in how they ask about LTC use and LTC insurance. For example, HRS, SHARE, and KLoSA ask about LTC insurance as part of their health services and insurance module, while ELSA asks about LTC insurance as part of their expectations module. The underlying difference reflects alternative

approaches: the first aimed at understanding coverage and use of private LTC insurance and the latter aimed at understanding how people are paying for LTC services or would intend to pay for such services should the need arise.

In developing a LTC utilization module section, it is important to understand a country's system for accessing LTC services. Understanding how a respondent may be covered for LTC services in the event they need them can help avoid asking inapplicable questions. Countries with universal LTC insurance may choose to ask more questions about expectations or use (e.g., KLoSA), while countries with limited coverage through public LTC insurance systems may need to ask more questions about private plans, their associated premiums, and covered services. We recommend HRS-INS surveys try to address at least the following questions about LTC service use:

4.3.1. ***Nursing home/residential care use***. Most HRS-INS ask about recent stays in nursing homes or residential care facilities but approach the topic in different ways. We recommend asking nursing home or residential care use for the past 12 months, whether these services were covered by insurance and how much the respondent paid out-of-pocket. Methodological issues: The HRS restricts to facilities providing 24-hour skilled nursing care and asks about stays since the last interview or in the last two years. SHARE asks about stays in any residential care facilities but asks additional questions to understand if there is at least one nurse on staff at the facility. The characterization of nursing home or residential care may be country-specific. Further, the HRS asks a respondent to report the number of nights or months, whereas SHARE asks respondents to report the number of weeks (rounding up – one day would be treated as one week). The HRS iterates over the reported number of nursing home stays and treats each as an episode, collecting information about the start and end of the episode and how the respondent paid for the nursing home stay during that episode.

4.3.2. ***In-home care use***. We recommend asking about in-home care services received, distinguishing between in-home medical care (e.g., care provided by a trained medical professional, such as nurse, physician assistant, nurse's aides, physical or occupational therapists, chemotherapists, respiratory oxygen therapists, or hospice caregivers) and care provided for personal needs (e.g., showering, dressing) or for daily tasks (e.g., cleaning, cooking). We also recommend asking about whether these services were covered by insurance and how much the respondent paid out-of-pocket.

5. **Family, Family Transfers, and Social Networks** Most HRS-INS interview the spouse of respondents regardless of the spouse's age, but several surveys do not interview the spouse and there are also cases with multiple spouses. Once spouses are interviewed, they are followed with longitudinal interviews, even if couples get divorced or separated.

5.1. ***Family composition***. All HRS-INS collect information about marital status and ask whether they have a living (1) mother and (2) father, and the numbers of living (3) sons, (4) daughters, (5) grandchildren, (6) brothers, and (7) sisters. Only a few studies ask about the birth order of respondents, which may be helpful in determining childhood (dis)advantage or pressure to provide care to elderly parents, or about deceased children and siblings, and biological (vs. step- or adopted) relationships.

5.2. ***Living arrangement***. All HRS-INS collect information on household size and whether the respondent co-resided with their spouse, children, parents, and/or others. This information is generally obtained in the household roster or coverscreen, which also collects a limited amount of socio-demographic information on the other household residents. Methodological issue: We have found that it is helpful to collect each household resident's relationship to the respondent and their spouse, if applicable, to ensure that correct relationships can be determined in later parts of the survey. Certain studies (LASI, ELSI) collect each resident's relationship to the head

of household, but if the head of household is not a respondent, it becomes impossible to accurately categorize relationships for the selected respondent(s). Further, collecting the residents' relationships to both respondent and spouse allows for more nuanced study, for example, which person's parent co-reside or co-residing biological versus step-children.

5.3. ***Family demographics***. Most HRS-INS collect information on age (current age or age at death), education, marital status, and living arrangements for parents and children. Age and education information for siblings are not often asked.

5.4. ***Family financial transfers***. All HRS-INS ask about financial transfers to and from the participant's children, parents, and other family members and friends. In addition to financial transfers, several HRS-INS (KLoSA, LASI, CHARLS) ask about in-kind transfers. We recommend asking about such in-kind transfers only in the local contexts where such transfers matter. <u>Methodological issues</u>. (1) Co-residence: Several surveys (ELSA, CHARLS, KLoSA) exclude transfers to and from co-residents, as resources tend to be shared within the household. However, transfers within the households contribute to individual income and impact relationships among household members and we therefore recommend capturing financial transfers within the household as well. (2) Censoring: Many surveys ask about transfers with some threshold amount (e.g., $500 or more for HRS), while other surveys ask about all transfers without a minimum transfer amount. (3) Periodicity: We recommend asking for the occurrence over the past 12 months. (4) Placement: In HRS and other high-income country surveys, financial transfer questions are included in the family module, whereas in several LMICs, these questions were asked in the income.

5.5. ***Care Provision***. We recommend placing questions asking about informal caregiving the respondent may provide in the Family module, as well as collecting the relationship with the care recipient, whether they assume primary caregiving responsibility or whether it is shared with someone else, and the frequency and duration of care, which will allow for the quantification of their caregiving burden. In addition, questions asking the respondent about providing care to grandchildren and persons with dementia can be useful.

5.6. ***Social interaction***. Most surveys ask about the frequency of contacts: (1) in-person and (2) via phone, mail, email or social media, with (1) parents, (2) children, (3) other relatives, and (4) friends. The relationships asked about are not always consistent across studies, neither are the answer scales for the frequency of contact. We recommend the following scale which was chosen to maximize comparability across existing surveys: 1. Daily or almost daily, 2. Several times a week 3. Once a week, 4. Every two weeks, 5. Once a month, 6. Less than once a month, 7. Almost never or never.

5.7. ***Social support***. Several surveys ask about social strain including the assessment of social relationship targets as critical, irritating, and unreliable. Assessments of emotional support availability, reliable alliance, and validation support also appear in measures of support quality across multiple surveys, assessing them with spouse, children, relatives, and friends.

5.8. ***Social and religious activities***. Most surveys ask about the frequency of engaging in social and religious activities, however the frequency of participation ranges from Likert scales, to weekly or even yearly. The most commonly named activities include: (1) volunteering or charity work; several cognitively stimulating activities, including (2) educational or training courses, (3) reading books, newspapers; (4) word or number games, puzzles; (5) playing cards, chess, and (6) using the computer for email, internet; religious and social activities, including (7) attending religious services, (8) participating in religious groups; (9) going to a sport, social, or other club, (10) participating in political or community-related organizations, and (11) going to a senior citizen's centers; as well as more solitary activities, including (12) home or car maintenance, gardening; and (13) watching TV.

6. Employment, Retirement, and Pension

6.1. ***Employment.*** All surveys include detailed questions to assess current labor force participation, job characteristics, and last employment if not currently working. In LMICs, the job market tends to be informal and characterizing informal job markets requires additional questions on second and third jobs, as well as consideration of different contractual arrangements.

We recommend surveys consider using a job loop that flexibly allows respondents to define how many jobs they currently have, including jobs related to being an employee, being self-employed, and doing agricultural work. As part of the job loop, a respondent identifies each unique job and then answers questions about the type of job, firm characteristics, supervisory responsibility, work hours and whether they are entitled to paid leave for sickness or vacation, pay, and the respondent's perceptions about their job, including job requirements, satisfaction, stress, and retirement timing. To reduce respondent burden, surveys can choose whether all jobs ask the same questions, or if certain questions are only asked of the respondent's primary job that requires the most work hours.

Outside of the job loop, questions are asked about on-the-job search for new employment, last employment if not currently working for pay, and unemployment and retirement. Key concepts to be captured in the employment module include:

6.1.1. ***Current employment status.*** Respondents must satisfy one or more of the following employment categories: (1) engaged in work for a salary or wage; (2) self-employed, or working for a family business/farm; (3) temporarily not working because of leave; (4) unemployed; (5) retired; (6) permanently sick or disabled; or (7) out of the labor force for other reasons. It is important that the survey clarifies if they are currently doing any paid work or have ever done paid work. Follow-up questions are asked of respondents who are currently employed or have ever been employed.

6.1.2. ***Job type and tenure.*** All jobs should be assigned one of three mutually exclusive job types – wage/salary jobs, self or family-employment in a non-agricultural sector, and self or family-employment in the agricultural sector, and job tenure should be collected. Methodological issue: In LMIC, additional questions may be asked to further characterize the informal nature of the job.

6.1.3. ***Job industry and occupation***: For at least the primary job, jobs should be classified using standardized occupation and industry codes. Multiple code frames have been used by various surveys, often following the already-established codes within a country (e.g., U.S. Census Occupation Code and Industry Code, International Standard Classification of Occupations (ISCO) and Nomenclature of Economic Activities (NACE) industry codes. For agricultural jobs, narrower occupational and work responsibilities may be asked. Methodological issue: It is important to note that code frames change over time within the country, and across countries, variations within the code frames are even greater. While it is useful to follow the established within-country code frames, for longitudinal analysis, it will be important to calibrate the code frames when the frames are updated. For the development of the recommended questionnaire, we have used ISCO and NACE codes, but surveys should decide whether these categories adequately map to their country-specific job categories, and if not, they should use country-specific categorizations.

6.1.4. ***Firm and workplace characteristics.*** For at least the primary job, surveys should clarify if a job is for a public or private employer, covered by a union, and the firm's overall size. These characteristics are often tied to a respondent's benefits (e.g., public pensions or insurance). The number of workers at the respondent's workplace is often collected as well.

6.1.5. ***Job supervisory responsibility.*** For at least the primary job, surveys should clarify if a respondent has a supervisory role in their job. Studies occasionally ask the number of employees the respondent is responsible for supervising.

6.1.6. ***Job hours and leave.*** For all recorded jobs, surveys collect total hours worked in a usual week and whether hours vary from week to week. Surveys will also often collect information on how many weeks an individual usually works in a job (this is important for seasonal jobs), and the amount of paid vacation or sick leave. For self-employed and business owners, it is important to indicate that the hours worked should include time doing books, tax, and or other duties that would require time.

6.1.7. ***Job pay***. For all recorded jobs, surveys collect information about earnings and hours of work, enabling an estimate of wage rate. The unit of time used to measure time worked and earnings (i.e., days, weeks, months, year) differs across the surveys, and we recommend the unit of time be determined by the respondent and standardization made after the survey or by the CAPI. Earnings and hours worked might include or exclude paid and/or unpaid overtime and meal breaks, so further instruction on inclusion or exclusion is needed. For self-employed and business owners, earnings should be referred to after taking out expenses rather than reporting revenue. <u>Methodological issue</u>: In agricultural settings, payment for work can be made in the form of goods rather than money.

6.1.8. ***Job requirements, perceptions, satisfaction, and stress***. HRS introduced a series of questions from Karasek's job demands and strain model (1979), which were later adopted by several surveys when asking about the respondent's primary job. The characteristics of a job, particularly whether it requires physical or cognitive demands have implications for health and cognition. Additionally, we identified six job stress questions that are most frequently asked in the HRS-INS regarding the primary job. We recommend these questions are asked of primary jobs.

6.1.9. ***Job continuity and retirement***. Work arrangements vary by country and some jobs are based on limited term contracts. For all jobs, the survey should clarify whether a job is permanent, a short-term contract, or temporary and for short-term or temporary work, an individual's expectation for its duration. Surveys may also ask specific questions about a respondent's expectation from retiring from their primary job and when people usually retire from a particular job.

6.1.10. ***Job search and unemployment***. Questions about whether a respondent is looking for a job are often asked for those who are not working for pay, and in a few studies for those who are currently working, with follow-up questions on the characteristics of a new job the respondent is looking for (e.g., part-time or full-time work). For unemployed respondents, questions in the HRS are aimed at eliciting the circumstances that separate an individual as unemployed versus not in the labor force. These may vary based on a country's definition of unemployment and surveys should balance accurately representing their country with cross-country comparability. In the US, the key threshold is based on active job search in the most recent 4 weeks (Bureau of Labor Statistics, 2024).

6.1.11. ***Last job worked***. For those who are not currently working, we recommend asking the same set of questions on job characteristics, earnings and wage rates, occupation and industry, and the primary reason for leaving the last job.

6.2. ***Retirement.*** Retirement is not a well-defined concept. For some individuals, they may think of retirement as leaving all work or, alternatively, as leaving their career job. A person continuing to work after stating they are retiring may view continued employment as a leisure activity. We recommend asking about a respondent's perceived retirement status, as this perception may inform their motivations for continued work or nonwork.

6.2.1. ***Retirement experience.*** We recommend questions capture: (1) if the respondent considers themselves to be retired, and if so, (2) when they consider themselves to have started retirement (either calendar year or age at retirement), (3) whether it is an early retirement, (4) the usual timing of retirement from the last job they worked at before they consider

themselves retired, (5) their main reason for retirement, and (6) their satisfaction with retirement.

6.2.2. **Retirement planning**. For those who have not retired, questions about retirement plans are asked, including when they plan to retire (calendar month year or age) and whether they plan to gradually reduce their work.

6.3. **Retirement Benefit Plans (Pensions and Retirement Savings Accounts)**

Information on a respondent's public and private pension plans and retirement savings accounts, collectively referred to as retirement benefit plans, are collected by most HRS-INS core surveys. Surveys typically ask what a respondent receives at the time of the interview or expects to receive in the future. This allows researchers to estimate pension wealth and predict future income from pensions. By doing so, researchers can analyze the effect of retirement plan design on retirement decisions and the relative contribution of retirement benefit plans to economic security in old age (Lee, 2010).

In developing a module for retirement benefit plans, it is important to understand a country's system for public and private retirement benefits and to identify the typical public and private pension and retirement savings plan designs. Key plan designs are: (1) plans that pay a regular benefit amount, (2) plans that accrue a balance that can be drawn on after benefit eligibility, and (3) plans that pay lump sum benefits (less common). A country can have multiple plan types. For example, in the US, Social Security is a retirement benefit plan that provides a defined benefit amount, while many private employers now provide their employees at least one retirement plan that accrues a balance. Understanding how a respondent is required or can save for retirement can help avoid asking inapplicable questions.

As many countries have retirement benefit plans tied to employers, a challenge that has arisen is that employer change can lead to dormant plans (i.e., plans that a respondent no longer contributes to but remains entitled to). These dormant plans can be overlooked. The HRS has moved to use a pension grid, where each retirement benefit plan is assigned a unique sub-respondent ID and is associated with the plan's true name and some details. This facilitates cross-interview consistency by ensuring returning respondents can recall and separately identify retirement plans. We encourage the use of a pension grid to track retirement benefit plans over time, including identifying account rollovers, closures, and dormancy.

After identifying each retirement benefit plan an individual owns or is entitled to as part of an initial screening, there are four general categories of information asked about each plan.

6.3.1. **Plan type**. Four questions identify key plan characteristics: (1) who is responsible for operating the plan? (2) does or will the plan pay a regular benefit? (3) does it accrue a balance? (4) Is the plan voluntary?

6.3.2. **Plan contributions.** For each plan, studies should ask the respondent's contribution amount or percentage of current salary, when the contributions started, and when they ended (if applicable). For plans that accrue a balance, studies should ask whether and how much the respondent's employer or the government contribute to the account, when those contributions started, and when they ended (if applicable).

6.3.3. **Plan eligibility.** Studies should ask individuals currently receiving benefits from a plan about the month-year when their benefits began or the age (including months) they started their benefits. For respondents who have not started receiving benefits from the plan, studies should ask when a respondent is first eligible to start benefits and when they anticipate starting benefits. It is important to ask about age in months or the month and year benefits began as many retirement plans have eligibility thresholds based on ages in months.

6.3.4. ***Plan benefits.*** Studies should ask about the amount current beneficiaries receive and the amount future beneficiaries anticipate receiving. For plans that accrue a balance, studies should ask the current balance and about how the balance is drawn down. For all plans, studies should ask about whether the benefit amount, once it begins is fixed or may change over time based on cost-of-living or economic conditions. Studies should also ask about separate plan entitlements to lump-sum payments and for the respondent's spouse and/or children (e.g., spouse and survivor benefits).

7. Wealth, Income, and Consumption

In developing an economic module, there are several important issues:

(1) *Economic unit and reporting.* In high income countries, economic data are often collected at either the individual- or couple-level, whereas in LMICs where household size is larger and older respondents are often economically dependent on adult children, the household is often considered as the appropriate economic unit. Whether the couple or the household are considered to be the economic unit, studies should select a financial respondent who is most knowledgeable about the finances of the unit who will provide answers to the economic modules on behalf of the entire economic unit.

(2) *Individual components versus total value.* In assessing the value of economic resources, it is better to ask for the value of individual components instead of asking for a total value. Asking respondents to sum across multiple components will result in less accurate economic reporting. We therefore recommend a set of broad components for wealth, income, and consumption in this paper that can be adapted to reflect the local context.

(3) *Unfolding bracket questions.* In answering economic value questions, respondents often are not able to or do not want to give specific values. In such cases, we recommend using a series of up to three unfolding bracket questions (Juster & Smith 1997). Each unfolding bracket question should require respondents to answer more than, about, or less than a specific value. This three-point scale is preferred to asking respondents to answer with a yes or no to a question of whether the value was above a specific value which can introduce confirmation bias. Across all three possible unfolding bracket questions in the set, we suggest using the following percentiles to determine the appropriate bracket values in the set of unfolding bracket questions: 7, 25, 50, 75, 93. At the point that a respondent has not provided a financial value, the survey would then start the set of unfolding bracket questions. For the first unfolding bracket question the initial bracket value can be selected randomly from the 2nd – 4th threshold value to mitigate any possible anchoring effects. The selection of the threshold of the subsequent two unfolding bracket question values will be determined by the initial random selection for the first bracket questions and respondent's response (see detailed skip pattern in Appendix Table). From the responses to up to three unfolding bracket questions, the value will fall into either one of five closed brackets (0 to 7th, 7th to 25th, 25th to 50th, 75th to 93rd percentile), approximate threshold values, or open bracket [93rd or higher].

(4) *Before or after tax.* Income payments may be reported before or after taxes and deductions, and it is important to be explicit about it. The lack of a uniform approach in treating taxation in the HRS-INS may be due to the way individuals think about their income in different environments and justified by the need to obtain more accurate country-specific measures, but it would be important and useful to provide information that will allow tax conversion for cross-country analysis.

7.1. ***Wealth***. Because of institutional differences in financial markets, different availability and spreading of investment products, as well as distinct saving habits, household wealth composition is very heterogeneous across countries. Consequently, despite the harmonization effort of the survey instruments discussed here, important measurement differences may inevitably persist across studies. To maximize meaningful cross-country comparisons, we make

the following broad asset groupings, whereas each individual country survey may want to itemize different asset types within each broad asset group. As income can be generated from assets, we recommend also asking about income generated from assets in the wealth module. Methodological issues: (1) While most HRS-INS consider the couple or household as the preferred economic unit, there are several surveys (e.g., KLoSA, LASI) aiming to capture individual-level asset ownership to study intrahousehold resource allocation. (2) Because some assets are expected to change slowly over time, additional questions can be introduced at the second wave of the study to verify large, reported value changes in specific assets between the current wave and the previous wave.

7.1.1. ***Primary residence and other real estate***. For the primary residence, we recommend first asking about: (1) ownership, followed by (2) the market value of the primary residence for homeowners, (3) the amount of the mortgage and home equity debts still owed on property, and (4) whether any rental income is generated, and if so, (5) the amount of income generated in the previous calendar year either before or after tax. A similar set of questions are recommended for other real estate and land properties, but noting to exclude any farm or business assets, which will be asked separately. Additional real estate and land questions can be looped so that the respondent reports separately about each property and does not require them to sum across multiple properties to calculate the total market value, total outstanding debt, or total rental income.

7.1.2. ***Housing characteristics***. We suggest collecting information about the housing characteristics of the primary residence. Despite a lack of consistency in the questions asked, more common questions across the HRS-INS include the number of rooms and bedrooms the household occupies, presence of disability-friendly household features, and fuels used for cooking or other household uses. Importantly for low-and-middle income countries, we recommend the inclusion of questions pertaining to access to a toilet, clean water supply, electricity, and cooking conditions.

7.1.3. ***Farm or business assets***. We suggest first asking about the ownership of any agricultural and non-agricultural businesses and their net values and then the amount they expect to receive if those farms or businesses were sold and they paid off any associated business debt.

7.1.4. ***Other non-financial assets.*** We suggest asking about a set of other non-financial assets by asset groupings, such as livestock, vehicles, and other equipment. It is important to remind the respondent to exclude any assets that they might have already reported as part of their business assets. For each asset grouping, respondents can be asked first about: (1) ownership of any asset in that asset grouping, followed by (2) the market value of across all assets they own in that grouping, (3) any financial liabilities associated with those assets and the amounts, and (4) any rental income from these assets. In LMICs, additional questions asking about both ownership and net value of consumer durable assets, such as refrigerators, TVs, etc. are common, whereas high-income countries rarely ask about these types of non-financial assets.

7.1.5. ***Financial assets.*** Because of differences in the way financial markets operate and in the availability of different investment products across countries, the list of financial investment products may vary across countries. Questions about the ownership and value of retirement accounts, bank accounts, stocks and mutual funds, and bonds are most frequently asked, but other financial products can be more common in some countries. For example, postal accounts may be a relatively more common saving vehicle. Private savings schemes in India (kitty parties, chit funds, and bishi), housing funds in China (Jizikuan), or private saving clubs in South Korea (Gye) may not have direct analogs in other social and economic contexts but are popular financial assets within a country. We suggest naming all popular financial products within a country. These questions can be asked by first: (1) asking about any ownership of this financial asset type, (2) the best estimated current value of the asset,

and (3) whether they received any interest income in the previous calendar year before or after tax and if so, how much they received. <u>Methodological issues</u>: (1) We also suggest asking about the value of any tax-advantaged or special-purpose retirement savings account in the pension module, though at the individual level, it would be important not to double-count these assets when creating a total income measure. However, the respondent reports from the pension module could be used to better understand the distribution of retirement accounts between those who are interviewed in the economic unit and those who are not, like a non-respondent spouse. (2) It is also possible that respondents may not know how to categorize assets across different asset types; for example, a tax-advantaged retirement account could be made up of stocks and mutual funds. It is important to remind respondents not to report the asset twice across different asset types, which would lead to double-counting, and for the study to provide clear instruction as to how each asset class is defined.

7.1.6. **Debts.** Like financial assets, debt vehicles may vary across countries. We suggest asking about several important categories of debt that might be important to distinguish between and making sure not to re-ask about debt that has already been collected earlier in the wealth module, like mortgage debts and other liabilities associated with real estate, farms, and businesses. We suggest the debt categories of credit card balances, educational debts, medical debts, personal loans from family and friends, and any other debts not asked about, but these should be amended depending on the country context.

7.2. **Income**. Like wealth, income sources are heterogeneous across countries. To enable meaningful cross-country comparisons, we propose the collection of information for six broad components of income: (1) business income, (2) earnings from wages or salaries, (3) pension and retirement account income, (4) public transfer income, (5) private transfers, and (6) other income. We have excluded asset income from the income module as this was already asked as part of the wealth module. Because we include a component of other income in the income module, this module must follow after the asset module so that once asked for any other income not already reported, the respondent does not report asset income, which would lead to double counting. We suggest using the previous calendar year as the reference period from which to collect income both for comparability to other HRS-INS studies and because it may allow respondents to refer to tax documentation in countries where the tax year aligns with the calendar year. While it may be reasonable to ask the respondent to report income over the entire previous year, for some other income components that are paid regularly and do not vary, for example, some public transfer programs, it may be better to allow respondents to report average monthly or quarterly income. If different reporting periodicities are allowed, it is important that all income measures are converted into annual amounts when creating total income measures. As economic resources are often pooled within a couple or household, it is important to collect information about the proper economic unit for the survey context.

7.2.1. **Business income**. Business income can capture income both from businesses that the couple or household owns and their self-employment. This can also include business and self-employment, which are agricultural like farms, fisheries, and forestry or non-agricultural. For couples or households who own multiple businesses, we suggest asking about each business separately to require less summation by the respondent. For each business, we suggest first asking: (1) the name of the business to allow for better recollection during follow-up interviews; (2) who owns or co-owns the business across all the members of the couple or household; (3) the best estimate of the total revenue from the business; (4) the best estimate of total costs of the business including rentals, materials, transportation, marketing, wages, taxes, and any other fees; (5) a confirmation of whether the respondent agrees that the profit for this business is the value of the reported revenue minus the reported costs, and if the respondent does not agree, the respondent is then asked to provide their best estimation of the profit from this business.

7.2.2. ***Earnings from wages and salaries.*** Earnings from wages and salaries can include income both from full-time employment and from odd jobs. It is important that respondents report the total value of their compensation, including bonuses, commissions, and overtime pay in the past calendar year. Because business income has already been asked, respondents should be reminded not to re-report any income here that they already reported as business income. To minimize the amount of summation required by the respondent and to better understand the intrahousehold income dynamics, we suggest looping through each member of the couple or household who receives wages/salaries. For each of these people, we ask what was the total amount of income they received in wages and salaries in the past calendar year before or after taxes. Methodological issue: We also suggest asking respondents questions about their jobs' pay rates as part of the employment module though the estimates of annual earnings income from those wage rates questions in the employment module may differ from these reports as the questions in the income module are designed to make sure to elicit compensation which may not be part of their regular pay like overtime, bonuses, tips, and commissions.

7.2.3. ***Pension and retirement account income.*** Retirement income is an especially important concept to capture to understand the level of income older people receive from plans and accounts which are designed to provide old-age income. This income component can include public, occupational, and private, pensions and retirement accounts. It can also include both pensions, which pay a regular amount based on a history meeting some eligibility like a required level of contributions, or tax-advantaged retirement accounts, which have an associated balance from which older people make withdrawals as a means of providing old-age income. To minimize the amount of summation required by the respondent and to better understand the intrahousehold income dynamics, we suggest looping through each member of the couple or household who receives pension or retirement account income. For each of these people, we ask for the total amount of income they received from pensions or retirement accounts in the past calendar year before or after taxes. Methodological issue: We also suggest asking respondents questions about their pension and retirement account income as part of the pension module. Questions in the pension module asks about current pension income, whereas questions in the income module asks about total income in the last calendar year. Further, the reports from the income module could be used to understand the retirement income of anyone in the economic unit who was not interviewed, like a non-respondent spouse.

7.2.4. ***Public transfer income.*** Like pension income, it is extremely important to capture the income respondents and their spouses or households receive from government assistance and support programs. The most common non-pension government transfer incomes include unemployment benefits, veteran's benefits, and welfare benefits; these will differ in naming and design across almost all countries. Most countries will also have other programs specific to their country. As respondents may recognize these benefits by program names, we suggest that studies list out all major public transfer programs by name through which people could receive income. In addition to cash transfer programs, it would also be important to collect information about programs with other forms of subsidy, such as a housing subsidy. This list should include programs which provide income to individuals and programs that provide income to households. To understand the level of income support each program provides, we suggest asking about each program that the respondent identifies anyone in the couple or household is receiving income. For each of these programs, we suggest asking first: (1) who in the couple or household has received income from the program with an option for the respondent to indicate that the income is jointly received by everyone in the couple or the household; and then (2) how much income from each program they receive in the past calendar year before or after taxes and deductions. Methodological issue: In our recommended questionnaire we have asked the respondent to mentally calculate a sum of the income across all people in the couple or household who receive income from a particular

public transfer program. For especially large households, this might be a difficult task and studies could also opt to ask the respondent to report the amount of income for each person in the household for each public transfer income, though this will increase the survey complexity and interviewing time.

7.2.5. *Private transfers*. Financial support from friends and family represents an important part of economic support for many older families around the world and especially in LMICs. We recommend asking about total transfer income, including inheritance and financial gifts, received by anyone in the couple or the household in the past calendar year before or after taxes. Methodological issues: We also suggest asking respondents questions about their individual private transfers from friends and family as part of the family module. It would be important not to double-count these two different reports of private transfers when creating a total income measure, but the respondent reports from the family module could be used to better understand private transfer income received at the individual level.

7.2.6. *Other income*. Other income is designed to capture any other sources from which the couple or household may have received income in the past calendar year that were not specifically asked about already in the survey. It is important for the question to highlight any specific income types that might be income sources but were not specifically asked about already. Commonly this includes lottery winnings, medical claims, life insurance payments, legal awards, alimony, dowry, and private scholarships. We suggest asking the respondent to report the total amount of other income received in the past calendar year across all members of the couple or household before or after taxes. Methodological issue: It is important that this question comes last across the collection of all wealth and income measures to avoid respondents reporting any income here which they would then be asked about later in the survey, which would result in double-counting.

7.3. *Consumption*. Expenditures and consumption are especially important when evaluating the resources utilized by a household. While "expenditures" is used to capture household purchases, "consumption" is often used as a broader term and incorporates the value of goods and services utilized by the household, including those not purchased, e.g., home-grown food. Consumption is most often collected in LMICs and at the household level, that is, the level of the unit that commonly shares resources and for whom it might be difficult to further separate out by individuals or couples. In HICs consumption is sometimes collected at the couple level. We recommend asking the value of consumption in reference to the period over which a household most often purchases or consumes that good or service, so that it requires less summing or averaging by the respondent. Methodological issues: (1) Developing a measure of total household consumption is complicated as it requires asking about all possible goods and services a household may consume over the course of 12 months. For a multi-purpose survey, we are limited with survey time and therefore refrain from recommending a comprehensive consumption survey but instead designed a consumption module that elicits measures of significantly important and internationally comparable consumption components. (2) For certain categories of consumption, it is expected that all couples or households would have consumed them. However, for other categories, some households may not have consumed these items at all in the 12 months, while others may have. For components where not all households have the specific consumption, it is important for surveys to either include an initial question asking if the household consumed the item or instruct respondents to enter '0' if they did not consume it.

7.3.1. *Food consumption*. We recommend asking about food consumption from either the past 7 days or the past 1 month, including the value of home-grown food and in-kind food transfers. We also recommend separating out food consumption inside and outside of the home. For both, the respondent is asked to estimate the total amount consumed across all couple or household members given the reference period.

7.3.2. ***Healthcare expenditures***. We recommend asking about several different types of health insurance instead of asking about healthcare expenditures as a whole. Specifically, our survey includes questions on: (1) health insurance premiums, (2) hospital visits and stays, (3) nursing home stays, (4) medication, (4) doctor's fees for outpatient visits, including both primary and specialist care, (5) outpatient survey, (6) medication, (7) dental care, and (8) any other medical expenses not already mentioned, like ambulances, medical equipment or in-home medical care. For each component, the respondent is asked to estimate the total expenditure amount across all couple or household members. We allow them to answer either with the average monthly amount over the past 12 months or the total amount in the past 12 months. While studies may choose to include unfolding brackets for each component, we have opted instead to include a question asking the respondent to estimate the total healthcare expenditures across all categories if they had not been able to provide estimates for each separate component of health expenditures. <u>Methodological issues</u>: (1) We also suggest asking respondents questions about their individual healthcare expenses as part of the healthcare utilization and insurance module. It would be important not to try to combine these two different reports of healthcare expenditures as it would result in double-counting, but the respondent reports from the healthcare module could be used to better understand the distribution of healthcare expenditures between those who are interviewed in the economic unit and those who are not, like a non-respondent spouse. (2) Whereas some surveys attempt to collect information about the expenses covered by insurance as well as out-of-pocket expenditures, respondents often report difficulties in estimating insurance coverage.

7.3.3. ***Recurring monthly expenditures***. Our survey attempts to collect a number of specific reoccurring expenditures that are common across many households and internationally comparable. These include: (1) communication fees; (2) utilities; (3) rent or mortgage payments; (4) transportation; (5) personal care, household help and other maintenance; (6) loan payments; (7) tobacco or other smoking products; and (8) entertainment. For each component, the respondent is asked to estimate the total expenditure amount across all couple or household members in the past 30 days. While studies may choose to include unfolding brackets for each component, we have opted instead to include a question asking the respondent to estimate the total amount the household spent on these recurring expenditures across all categories if they had not been able to provide estimates for each separate expenditure component.

7.3.4. ***Other expenditures***. Our survey attempts to collect a number of additional expenditures that are common across many households and are internationally comparable. These include: (1) clothing and apparel; (2) education and training; (3) durable goods; (4) insurance payments (excluding already reported healthcare insurance premiums); (5) expenses related to religion or community groups; (7) vacation, hobbies, jewelry, or other luxury items; (8) home improvements or car repairs; or (9) any other expenses. For each component, the respondent is asked to estimate the total expenditure amount across all couple or household members in the past 12 months. While studies may choose to include unfolding brackets for each component, we have opted instead to include a question asking the respondent to estimate the total amount the household spent on these other expenditures across all categories if they had not been able to provide estimates for each separate expenditure component.

8. Childhood Experiences

Childhood experiences may have a lasting influence on adult health and cognition, and several surveys attempted to capture this information either through life-history interviews or leave-behind or mailed questionnaires. We suggest a limited set of questions to capture childhood experiences, capturing: (1) socioeconomic status, (2) health conditions, and (3) home environment.
<u>Methodological issue</u>: When asking about childhood, some surveys specify the age as 10 or 16, other surveys ask more vaguely without specifying an age.

8.1. ***Socioeconomic status***. To measure crowding, we suggest asking the number of bedrooms and household size. We also suggest other housing characteristics, such as the availability of hot or cold running water, a toilet, central heating, electricity, as well as the availability of books, whether they grew up in a rural area, and their family's financial status.

8.2. ***Health***. We suggest assessing overall health status, whether they missed school due to a health problem, and their experience with specific illnesses.

8.3. ***Home environment***. A series of questions are asked about the home environment, such as parental alcohol or drug use, physical abuse, parental divorce, and separation from parents over 6 months.

9. Stress

The NIA Stress Measurement Network's leadership identified domains of stressful life experiences that are linked to health-related outcomes and proposed a set of stress domains and harmonized measures for cross-country analysis in partnership with the Gateway team (Gruenewald et al., 2020).

9.1. ***Lifetime trauma.*** We recommend Krause's 7-item lifetime trauma questions that have been used in several surveys, including HRS, to maximize the harmonization potential.

9.2. ***Loneliness.*** All surveys included at least one question about loneliness. We recommend the 3-item UCLA loneliness scale for inclusion.

9.3. ***Neighborhood disorder and lack of cohesion.*** We recommend Cagney et al.'s (2009) 7-item neighborhood physical disorder and social cohesion, as this scale has desirable psychometric properties and has been successfully used in HRS and ELSA.

9.4. ***Discrimination.*** Several surveys adopted the 6-item scale assessing perceived everyday discrimination (Williams et al. 1997) and reasons attributed for discrimination (Kessler et al. 1999).

10. Psychosocial Measures

The NIA Emotional Well-being Network and the Gateway team reviewed and grouped subjective well-being into four categories: evaluative, hedonic, eudaimonic, and experienced well-being (Smith et al., 2022). As most HRS-INS only include evaluative and eudaimonic well-being measures, our suggestion is limited to those.

10.1. ***Life satisfaction.*** We suggest Diener's 5-item life-satisfaction questions. Several surveys also include domain-specific life-satisfaction questions and an additional overall life-satisfaction question, which are also included in the Gateway questionnaire, however, a variety of scales have been used across existing HRS-INS surveys (Campbell, Converse & Rodgers, 1976).

10.2. ***Psychological well-being.*** We suggest inclusion of the 12-item CASP measure of psychological well-being.

10.3. ***Subjective social status.*** Several studies also included Cantril's ladder to measure how respondents perceive their social status (Cantril, 1965).

10.4. ***Expectations.*** As subjective expectations are likely to play an important role in intertemporal decisions, several surveys ask about the respondent's expectations about their own individual well-being and life events such as: (1) survival, (2) inheritance, and (3) retirement. Methodological issue: Methods of elicitation and visual aids vary across surveys. HRS, ELSA, SHARE, and JSTAR elicit a continuous measure from 0 to 100. KLoSA asks for response on a 10-point Likert scale, while LASI employs a similar concept using 10 stones that are physically counted. ELSA, SHARE and KLoSA also present visual aids of their scales. In JSTAR, respondents are shown a chart of average life-expectancies for Japanese men and women, but no visual aid of the probability scales to aid in giving responses. After an investigation of older individuals' subjective survival expectations in LMICs, Delavande, Lee and Menon (2017)

concluded that although individuals are, on average, able to understand the concept of probability, responses are sensitive to framing effects and own versus hypothetical person effects, calling for careful pretests and further investigation within the study context.

11. End of Life Planning

Though rarely included throughout the HRS-INS, factors concerning end-of-life are growing increasingly important and will be an important facet of study in the future.

11.1. ***Wills and trusts.*** Though highly context dependent, we recommend the use of questions pertaining to the existence of an official will or trust, as well as who is to receive assets in the event of death.

11.2. ***End of life wishes.*** We recommend questions asking whether the respondent has discussed preferred medical care at the end-of-life with anybody, the existence of a durable power of attorney for healthcare and the designated person, the existence of a living will, the respondent's desired level of care at the end-of-life, as well as their preference for death location and what factors are important in determining that location.

12. Interviewer Observations

We provided questions for the interviewer to answer at the conclusion of the interview.

12.1. ***Respondent observations.*** These questions ask the interviewer to identify anyone who helped the respondent answer questions and how much they answered, as well as any places in the interview that were difficult or bothersome for the respondent, and their overall performance with the interview.

12.2. ***Housing observations.*** These questions ask the interviewer to identify the location and type of home, the quality of its construction, and its accessible features.

13. Recommendations for HCAP Studies.

The HCAP is part of an international research collaboration funded by the National Institute on Aging to better measure and identify cognitive impairment and dementia in representative population-based samples of older adults, in the context of ongoing longitudinal studies of aging in high-, middle-, and low-income countries around the world. These data resources aim to expand research opportunities to study key life-course factors that likely affect cognitive function and the risk for Alzheimer's Disease (AD) and Alzheimer's Disease Related Dementia as well as the associated costs. Whereas the HCAP was developed as an add-on study to the HRS-INS, with the success of the HCAP studies, there has been increasing interest throughout the scientific community to develop new HCAP studies in the absence of ongoing longitudinal studies of aging. For these new HCAP studies without parent longitudinal studies, we suggest the inclusion of the following survey content from the core questionnaire to enable the investigation of key risk factors for and the cost of cognitive impairment and dementia. Our recommendation is based on the recent Lancet Commission report (2024) on dementia prevention, intervention, and care.

13.1.1. ***Risk factors for dementia.*** The 2024 update of the Lancet Commission identified 14 risk factors: education, hearing loss, high LDL cholesterol, depression, traumatic brain injury, physical inactivity, diabetes, smoking, hypertension, obesity, excessive alcohol consumption, social isolation, air pollution, and visual loss. The evidence suggests that it is important to be cognitively, physically, and socially active throughout life, and therefore, to better understand life-long risk factors, longitudinal data are essential. For studies without such data, retrospective interviews can provide most of the critical information.

In addition, we recommend collecting geographic information that will allow linkage to ambient air pollution data and other exposures associated with the physical environment. The Gateway to Global Aging Data team has curated various exposome data, including air pollution, and urge the HRS-INS teams to link such data to individual-level survey data using geographic information.

13.1.2. ***Hypothesized confounders for risk factor analyses.*** To facilitate the study of key risk factors for dementia using observational research methods, it is also important to collect information on potential confounders of risk factor associations as well. To ensure that the data allow for adequate confounder control in risk factor analyses, we recommend the additional collection of data on: demographic and socio-economic characteristics in adulthood (social class, race/ethnicity, etc.), economic status (wealth in high income countries, per capita income in LMICs (Angrisani, Lee, Rebellato, 2023)), childhood health and socio-economic status, and cardiovascular disease history.

13.1.3. ***Cost of dementia.*** A growing body of literature has estimated the cost of dementia, and many papers apply similar definitions of cost, that is, medical and long-term care costs and informal care (Hurd et al., 2013; Cantarero-Prieto et al. 2020, Meijer and Lee, 2024). With the goal of increasing comparability across the studies, we make the following recommendations based on this approach, leaving out intangible costs to the persons with dementia, such as pain, suffering and caregiver burden (El-Hayek et al., 2019).

In quantifying the cost, medical and long-term care costs are often split by payers, household out-of-pocket, insurance and government. A household survey usually does not provide reliable information about the amounts directly paid by third parties, such as insurance companies and government programs, whereas medical claims data may not cover costs paid by individuals who did not reach their deductible or costs paid by other third parties.

For valuing informal care, two main approaches have been frequently adopted (Engel et al., 2021). Replacement cost asks what it would have cost if the same hours of care were provided by a formal caregiver. Opportunity cost values hours of care at the wage rate of an informal caregiver who would earn if they spent those same hours working in a paid job. Pros and cons of these approaches are further discussed in Meijer and Lee (2024), but regardless of the approaches taken, to estimate the cost of dementia, we suggest including questions on informal care received and provided.

References

Angrisani M, Lee J, G Rebellato. "Consumption, saving and wealth at old age: Comparing evidence from developed and developing countries," In Sousa-Poza A, Bloom DE, Uwe S (Eds.), the *Routledge Handbook on the Economics of Ageing*, 2023, Routledge, ISBN 9780367713324

Bureau of Labor Statistics. 2015. Labor force statistics from the Current Population Survey, https://www.bls.gov/cps/cps_htgm.htm, Accessed on September 5, 2024.

Cagney KA, Glass TA, Skarupski KA, Barnes LL et al. 2009. Neighborhood-level cohesion and disorder: measurement and validation in two older adult urban populations. Journal of Gerontology: Series B, 64(3), 415-424.

Campbell A, Converse PE, Rodgers W. 1976. The quality of American life: Perceptions, evaluations, and satisfactions. New York: Russell Sage Foundation.

Cantarero-Prieto, D., Lanza Leon, P., Blazquez-Fernandez, C., Sanchez Juan, P., & Sarabia Cobo, C. (2020). The economic cost of dementia: A systematic review. Dementia, 19, 2637–2657.

Cantril H. 1965. The pattern of human concerns. New Brunswick, NJ. Rutgers University Press.

CDC, Centers for Disease Control and Prevention. 2000. Measuring Healthy Days: Population Assessment of Health-Related Quality of Life (HRQOL). https://archive.cdc.gov/#/details?url=https://www.cdc.gov/hrqol/pdfs/mhd.pdf. Accessed on July 22, 2024.

Delavande A, Lee J, Menon S. "Eliciting survival expectations of the elderly in low-income countries: Evidence from India," *Demography*, 2017, 54(2):673-699. doi:10.1007/s13524-017-0560-8. PMID:28281273

Domingue BW, McCammon RJ, West BT, Langa KM, Weir DR, Faul J. The Mode Effect of Web-Based Surveying on the 2018 U.S. Health and Retirement Study Measure of Cognitive Functioning. The Journals of Gerontology: Series B. 2023 Sep 1;78(9):1466–73.

Ehrlich JR, Agarwal A, Young C, Lee J, Bloom DE. 2022. The prevalence of vision impairment and blindness among older adults in India: findings from the Longitudinal Ageing Study in India," *Nature Aging*, DOI:10.1038/s43587-022-00298-6.

El-Hayek, Y.H., Wiley, R.E., Khoury, C.P., Daya, R.P., Ballard, C., Evans, A.R., Karrand, M., Molinuevo, J.L., Norton, M., & Atri, A. (2019). Tip of the iceberg: Assessing the global socioeconomic costs of Alzheimer's disease and related dementias and strategic implications for stakeholders. Journal of Alzheimer's Disease, 70, 323–341.

Engel, L., Ajdukovic, M., Bucholc, J., & McCaffrey, N. (2021). Valuation of Informal Care Provided to People Living With Dementia: A Systematic Literature Review. Value in Health, 24, 1863–1870.

Gross AL, Li C, Briceño EM, Rentería MA, Jones RN, Langa KM, et al. Harmonisation of later-life cognitive function across national contexts: results from the Harmonized Cognitive Assessment Protocols. The Lancet Healthy Longevity. 2023 Oct 1;4(10):e573–83.

Gruenewald TL, Crosswell AD, Epel E, Mayer S, Smith J, Lee J. 2020. Measurement of stress in the Health and Retirement Study (HRS) and the HRS Family of Studies, https://g2aging.org/cms/download/user-guide/Stress_Measurement_in_the_HRS_Family_of_Studies_Guide_5_14_20-1.pdf

Hurd, M.D., Martorell, P., Delavande, A., Mullen, K.J., & Langa, K.M. (2013). Monetary Costs of Dementia in the United States. New England Journal of Medicine, 368, 1326-1334.

Jorm AF, Jacomb PA. The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): socio-demographic correlates, reliability, validity and some norms. *Psychol Med*. 1989;19(4):1015-1022. doi:10.1017/s0033291700005742

Juster, F.T. and Smith, J.P., 1997. Improving the Quality of Economic Data: Lessons from the HRS and AHEAD. *Journal of the American Statistical Association*, *92*(440), pp.1268-1278.

Kapteyn A, Banks J, Hamer M, Smith JP et al. 2018. "What they say and what they do: comparing physical activity across the USA, England and the Netherlands, *J of Epidemiology and Community Health*, 72(6): 471-476.

Kapteyn A, Smith J, van Soest A. 2007. "Vignettes and self-reports of work disability in the United States and the Netherlands," *American Economic Review*, 97 (1): 461-473.

Karasek Jr. RA. 1979. Job demands, job decision latitude, and mental strain: Implications for job redesign, Administrative Science Quarterly, 24 (2): 285–308.

Kessler RC, Mickelson KD & Williams DR. 1999. The prevalence, distribution, and mental health correlates of perceived discrimination in the United States, Journal of Health and Social Behavior, 40 (3): 208-230.

Kobayashi LC, Gross AL, Gibbons LE, Tommet D, Sanders RE, Choi SE, et al. You Say Tomato, I Say Radish: Can Brief Cognitive Assessments in the U.S. Health Retirement Study Be Harmonized With Its International Partner Studies? The Journals of Gerontology: Series B. 2021 Nov 1;76(9):1767–76.

Krause N, Shaw BA, Cairney J. 2004. "A descriptive epidemiology of lifetimetrauma and the physical health status of older adults," *Psychology and Aging*, 19 (4): 637-648.

Lee J. 2010. "Data sets on pensions and health: Data collection and sharing for policy design," *International Social Security Review*, 63 (3-4):197-222.

Lee J, Phillips D, Wilkens J, Gateway to Global Aging Data Team. "Gateway to Global Aging Data: Resources for cross national comparisons of family, social environment, and healthy aging," *Journal of Gerontology: Social Sciences*, 2021, 76 (S1): S5-S16. doi:10.1093/geronb/gba050.

Livingston G, Huntley J, Sommerlad A, Ames D et al. 2020. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission, *Lancet*, 396 (10248): 413-446.

Livingston G, Huntley J, Liu KY, Costafreda SG et al. 2024. Dementia prevention, intervention, and care: 2024 report of the Lancet Commission, Lancet, 404: 572-628.

Meijer E, Lee J. 2024. "How should we measure the cost of dementia," University of Southern California, mimemo.

National Institute on Aging, National Institutes of Health. 2006. *The Health & Retirement Study: Growing older in America*. Data Book, Bethesda, MD.

Nichols E, Ng DK, Hayat S, Langa KM, Lee J, Steptoe A, et al. Differences in the measurement of cognition for the assessment of dementia across geographic contexts: Recommendations for cross-national research. Alzheimers Dement. 2022 Jul 16;

Nichols E, Gross AL, Zhang YS, et al. Considerations for the use of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) in cross-country comparisons of cognitive aging and dementia. *Alzheimer's Dement*. 2024; 20: 4635–4648. https://doi.org/10.1002/alz.13895

Nichols E, Jones RN, Gross AL, Hayat S, Zaninotto P, Lee J. Development and assessment of analytic methods to improve the measurement of cognition in longitudinal studies of aging through the use of substudies with comprehensive neuropsychological testing. *Alzheimer's Dement*. 2024; 1-13. https://doi.org/10.1002/alz.14175

O'Brien CP. 2008. "The CAGE questionnaire for detection of alcoholism," *JAMA*, 300(17): 2054-2056. DOI:10.1001/jama.2008.570.

Radloff, LS. 1977. "The CES-D scale: A slef-report depression scale for research in the general population." Applied Psychological Measurement, 1(3): 385-401.

Russell DW. 1996. UCLA loneliness scale (Version 3): Reliability, validity, and factor structure. *Journal of Personality Assessment*, 66 (1): 20-40.

Smith JR, Gibbons LE, Crane PK, Mungas DM, Glymour MM, Manly JJ, et al. Shifting of Cognitive Assessments Between Face-to-Face and Telephone Administration: Measurement Considerations. The Journals of Gerontology: Series B. 2023 Feb 1;78(2):191–200.

Smith J, Ryan L, Park Y, Crosswell A, Phillips D, Lee J. 2022. "User Guide: Subjective Well-being," CESR-SCHAEFFER Working Paper Series No. 2022-006. https://cesr.usc.edu/documents/WP_2022_006.pdf.

Suzman R, Harper S. 2013. "Editorial: Dawning of a new age for longitudinal cohort data," *Population Ageing*, 6:1-4. DOI:10.1007/s12062-013-9083-2.

Wang Y, Wilkens J, Phillips D. 2023. *User Guide: Physical and Anthropometric Measurement*, CESR-Schaeffer Working Paper Series No. 2023-008. https://cesr.usc.edu/documents/WP_2023_008.pdf

Williams DR, Yu Y, Jackson JS & Anderson NB. 1997. Racial differences in physical and mental health: socio-economic status, stress and discrimination. Journal of Health Psychology, 2: 335-351.

Appendix Table. **Skip pattern for unfolding bracket questions**

Threshold for the 1st questions	Response to the 1st question	Threshold for the 2nd question	Response to the 2nd question	Threshold for the 3rd question	Response to the 3rd question	Range identified
25th percentile	More than	75th percentile	More than	93th percentile	More than	93rd percentile+
					Equal to	93rd percentile
					Less than	75th – 93rd percentile
			Equal to	Skipped		75th percentile
			Less than	50th percentile	More than	50th – 75th percentile
					Equal to	50th percentile
					Less than	25th – 50th percentile
	Equal to	Skipped				25th percentile
	Less than	7th percentile	More than	Skipped		7th – 25th percentile
			Equal to	Skipped		7th percentile
			Less than	Skipped		0 – 7th percentile
50th percentile	More than	75th percentile	More than	93rd percentile	More than	93rd percentile+
					Equal to	93rd percentile
					Less than	75th – 93rd percentile
			Equal to	Skipped		75th percentile
			Less than	Skipped		50th -75th percentile
	Equal to	Skipped				50th percentile
	Less than	25th percentile	More than	Skipped		25th – 50th percentile
			Equal to	Skipped		25th percentile
			Less than	7th percentile	More than	7th – 25th percentile
					Equal to	7th percentile
					Less than	0 – 7th percentile
75th percentile	More than	93rd percentile	More than	Skipped		93rd percentile+
			Equal to	Skipped		93rd percentile
			Less than	Skipped		75th – 93rd percentile
	Equal to	Skipped				75th percentile
	Less than	25th percentile	more than	50th percentile	More than	50th -75th percentile
					Equal to	50th percentile

				Less than	$25^{th} - 50^{th}$ percentile
		Equal to	Skipped		25th percentile
		Less than	7th percentile	More than	$7^{th} - 25^{th}$ percentile
				Equal to	7th percentile
				Less than	$0 - 7^{th}$ percentile