



Data Resource Profile

Data Resource Profile: Cross-national and cross-study sociodemographic and health-related harmonized domains from SAGE plus ELSA, HRS and SHARE (SAGE+, Wave 1)

Nadia Minicuci,^{1,*} Nirmala Naidoo,^{2,3} Somnath Chatterji^{2,3} and Paul Kowal^{2,4}

¹National Research Council, Neuroscience Institute, Padova, Italy, ²SAGE, ³IER/HIS, World Health Organization, Geneva, Switzerland and ⁴Research Centre for Generational Health and Ageing, University of Newcastle, Callaghan, NSW, Australia

*Corresponding author. National Research Council, Neuroscience Institute, Via Giustiniani, 2 35128 Padova (Italy). E-mail: nadia.minicuci@unipd.it

Accepted 1 June 2016

Abstract

Four longitudinal studies were included in this rigorous harmonization process: the Study on global AGEing and adult health (SAGE); English Longitudinal Study on Ageing (ELSA); US Health and Retirement Study (HRS); and Survey of Health, Ageing and Retirement in Europe (SHARE). An ex-post harmonized process was applied to nine health-related thematic domains (socio-demographic and economic, health states, overall self-report of health and mental state, health examinations, physical and mental performance tests, risk factors, chronic conditions, social network and subjective well-being) for data from the 2004 wave of each study. Large samples of adults aged 50 years and older were available from each study: SAGE, $n = 18\,886$; ELSA, $n = 9\,181$; HRS, $n = 19\,303$; and SHARE, $n = 29\,917$. The microdata, along with further details about the harmonization process and all metadata, are available through the World Health Organization (WHO) data archive at [<http://apps.who.int/healthinfo/systems/surveydata/index.php/catalog>]. Further information and enquiries can be made to [sagesurvey@who.int] or the corresponding author. The data resource will continue to be updated with data across additional waves of these surveys and new waves.

Data resource basics

SAGE-plus (SAGE+) is part of the World Health Organization's Study on global AGEing and adult health (SAGE) effort to improve measurement strategies and comparability across large ageing studies, including harmonization of SAGE plus the English Longitudinal Study on

Ageing (ELSA), US Health and Retirement Study (HRS) and Survey of Health, Ageing and Retirement in Europe (SHARE). Two waves from each of these population-based longitudinal studies were harmonized to leverage the benefits of cross-national and cross-study comparisons that enable us to better understand the nature of demographic

and epidemiological transitions.^{1,2} This paper describes the process and resulting harmonized dataset for the baseline wave (SAGE+ Wave 1) of each contributing study. The main outcome is a common dataset with all corresponding metadata. The aim of this article is to provide details about the process and methods used for generating a new data resource on older adults and ageing in different regions of the world from four major data collection efforts.

Dataset production

Although harmonization, not homogeneity, is the goal, each thematic area requires attention to study standards, data quality, bias and measurement issues. This harmonization process borrows from the best practices developed by other harmonization efforts employed by collaborations such as the Cross-national Determinants of Quality of Life and Health Services for the Elderly (CLESA), International Household Survey Network (IHSN), Dynamic Analyses to Optimise Ageing (DYNOPTA), and Phenotypes and eXposures (PhenX).³⁻⁶ This was supplemented by regular communication with ongoing surveys and surveillance activities to ensure that sharing standards, data, tools and methodologies continues to inform the outcomes of this undertaking. Where possible, variables were harmonized to enable comparison with WHO benchmarks, such as physical activity variables that were harmonized to derive measures reflecting WHO recommendations for weekly physical activity targets, and alcohol consumption data that provided classifications in accordance with WHO recommendations.^{7,8}

To this end, the goal of the SAGE+ harmonization process was to generate a public-access dataset for assessing the associations between health and well-being in older ages and the ageing process across different individual, cultural, societal and policy backgrounds.

Data resource area and population coverage

SAGE consists of standardized longitudinal panel studies in China, Ghana, India, Mexico, the Russian Federation and South Africa [www.who.int/healthinfo/sage]. The survey is designed to be nationally representative of the population aged 50 years and older, with a smaller sample of adults aged 18-49 years, focusing on health and health-related outcomes and their determinants with an explicit aim to enhance cross-population comparability especially in lower- and middle-income countries. SAGE Wave 0 (2003/04) and Wave 1 (2008/10) were considered for harmonization. Wave 0 individual response rate ranged from 89% in South Africa to 100% in China.⁹

ELSA is a representative sample of the older English population, which began in 2002 with six waves of data

collection completed through 2015 [www.elsa-project.ac.uk]. ELSA was modelled on the US HRS, and explored a wide range of health and wealth/retirement domains.^{10,11} Datasets from Wave 2 (2004) and Wave 5 (2010) were considered for harmonization. Wave 2 individual response rate was 82.0%.

The US HRS is a landmark study on ageing combining socioeconomic, health and retirement issues over numerous waves, while pioneering many improvements in measurement methodology [hrsonline.isr.umich.edu]. HRS interviews more than 22 000 Americans aged 50+ every two years (baseline 1992), covering an array of topics including health, insurance, financial and labour market status, and retirement planning.^{12,13} HRS Wave 7 (2004) and Wave 10 (2010) were considered for harmonization. Wave 7 response rate was 86.4%.

The 2004 SHARE baseline consisted of cross-national panel surveys that collected health, socioeconomic status, social and family networks, labour market status, and retirement planning data on individuals aged 50+ [www.share-project.org] in 11 countries (Austria, Belgium, Denmark, France, Germany, Greece, Italy, Switzerland, The Netherlands, Spain and Sweden).¹⁴ Wave 1 (2004) and Wave 4 (2010) were considered for harmonization. Wave 1 individual response rates ranged from 74% in Spain to 93% in France.

In this paper, the harmonized variables referring only to SAGE+ Wave 1 are presented (SAGE, Wave 0; ELSA, Wave 2; HRS, Wave 7; SHARE, Wave 1), and consider 2004 as the SAGE+ baseline year. A future paper will present SAGE+ Wave 2 (2010).

Data collected and harmonization process

The process of harmonizing variables across this range of surveys built upon established procedures. The first step included documenting the study design and variables collected for each study; the second step determined the harmonizable domains and the group of core variables targeted for available domains; the third step addressed data processing where harmonizable variables were constructed either with precisely common variables that made the pooling quite straightforward or with heterogeneous variables that required decisions for their definition and categorization; and the last step assessed data quality using statistical indicators. A detailed documentation of the harmonization process was meticulously documented in the STATA code used to create derived variables across datasets, and will be distributed along with all metadata. This documentation meets the Data Documentation Initiative (DDI) standards and the international standards for data and metadata exchange.⁴

Harmonized variables were categorized into nine domains as follows.

- i. Socio-demographic and economic: sex, age, marital status, living arrangements, educational level, occupation, age at retirement.
- ii. Health states describing physical functioning: activities of daily living (ADL), instrumental activities of daily living (IADL), mobility, near and distance vision, hearing, pain.
- iii. Overall self-report of health and mental state: self-reported health, cognition/memory, depression, sleep, loneliness.
- iv. Health examinations: measured blood pressure.
- v. Physical and mental performance tests: normal-pace walking test, rapid-pace walking test, grip strength, cognition/memory tests.
- vi. Risk factors: body mass index (BMI), tobacco consumption, alcohol consumption, physical activity.
- vii. Chronic conditions: self-reported chronic conditions, disease treatment.
- viii. Social network: social network index.
- ix. Subjective well-being: quality of life, life satisfaction, well-being.

Socio-demographic and economic domains

The harmonization of de-identified socio-demographic and economic variables was quite straightforward across all four studies. The original education variables were recorded as the highest educational qualification obtained, the age at which school was finished and the number of years of education completed. The harmonized education variable was generated converting the qualification level into the number of years.

Physical functioning domain

[Supplementary Table A](#) (available as [Supplementary data](#) at *IJE* online) consists of a checklist of the physical functioning items from each study. The harmonization process was straightforward for ELSA, HRS and SHARE since the original items were phrased similarly. This included activities of daily living (ADL) and instrumental activities of daily living (IADL) questions. The SAGE variable was based on two self-care questions. Reporting no difficulty in any ADL/IADL resulted in being classified as independent; reporting difficulty with one or two ADLs/IADLs was considered a mild disability; reporting difficulty with three or four items represented moderate disability; and, difficulties with five or more items was classified as severe disability/cannot do.

Self-reported health, and physical and mental health domains

[Supplementary Table B](#) (available as [Supplementary data](#) at *IJE* online) shows the variables available for these domains by study. For overall general self-rated health (SRH), the single question in ELSA, HRS and SHARE all provided the same five response categories, from excellent to poor. SHARE also, like SAGE, used response categories from very good to very bad.

For cognition, ELSA and HRS asked respondents to rate their memory (from excellent to poor) whereas SAGE asked about difficulties in remembering (from none to extreme). The harmonized item reports the level of difficulty, where respondents who rated their memory as excellent were categorized as having no difficulties in remembering things.

For the affective component of mental health, self-reported depression and sadness, based on a single question about feeling depressed or sad, were evaluated separately in ELSA and HR, and conjointly in SAGE and SHARE. The harmonized variable was dichotomized into yes/no since only SAGE asked to grade the level of feeling sad or depressed (on a 5-point scale from none to extreme).

Self-reported trouble with sleeping was recorded in a heterogeneous way across surveys and only a dichotomous variable (yes/no) was possible for harmonization. Reported feelings of loneliness were available only for ELSA and HRS.

Health examination

Blood pressure measurements were available only for ELSA and HRS. The average of the second and third measurements was taken to classify hypertensive (mean systolic blood pressure ≥ 140 mmHg or diastolic blood pressure ≥ 90 mmHg or currently taking anti-hypertensive medications) and normotensive (<140 mmHg and ≤ 90 mmHg, and had not been taking anti-hypertensive medications during the past 12 months) subjects. Hypertensive respondents were classified as: hypertensive not on treatment; normotensive on treatment; or hypertensive on treatment.

Physical and mental performance tests domain

Walking test eligibility criteria differed across surveys; ELSA administered the test only to individuals aged 60+ years; HRS to those aged 65+ years; and SHARE to those aged 75+ years. SAGE Wave 0 did not administer a walking test. The same distance (2.5 m) was covered in each study, but had high missing, refusal or not-applicable

values: ELSA (valid $n = 5477$ from $n = 6149$); HRS (valid $n = 1864$ from $n = 11097$); and SHARE (valid $n = 2757$ from $n = 5534$).

Supplementary Table C (available as **Supplementary data** at *IJE* online) includes descriptions of the cognition tests used in each study. The tests consisted of four questions on time orientation (day of the month, month, year, day of the week) and two questions for immediate and delayed recall. HRS administered the time orientation test only to participants aged 65+, whereas ELSA and SHARE did not apply this age criterion.

Risk factors domain

Tobacco consumption: two different harmonized variables were constructed. One variable was based on three studies (ELSA, HRS and SHARE) and classified smoking status as never, current and not current. The other variable, based on all four studies, was constructed as dichotomous (currently or not currently smoking) because SAGE asked only about current smoking status. Mean number of cigarettes per day was harmonized for all four studies, and SAGE, HRS, and SHARE also had number of years smoking.

Alcohol consumption: abstainer status was generated for all studies. Among the non-abstainer group, heavy drinker was constructed as: 3+ days per week/past 7 days (in the past 3 months for HRS) with five/four or more standard drinks for men/women. Similarly, infrequent heavy drinkers were classified as 1 or 2 days per week/past 7 days (in the past 3 months for HRS) with five/four or more standard drinks for men/women. Days per week drinking alcoholic beverages and drinks per day were not available in SHARE.

Physical activity: both vigorous and moderate/mild was harmonized across all studies. Sedentary behaviour was also constructed for all studies, defined as: vigorous, moderate or mild physical activity performed hardly ever or never for ELSA, HRS and SHARE, and not performed in the past seven days for SAGE.

Chronic conditions domain

Supplementary Table D (available as **Supplementary data** at *IJE* online) lists the set of self-reported chronic health conditions recorded in each study through the question 'Have you ever been diagnosed with/told by a doctor that you have [...]?' The conditions available in at least three studies include: angina, arthritis, asthma, cancer, chronic lung disease, diabetes, hypertension and stroke. Current treatment for the harmonized conditions was also recorded, except for cancer where treatment in the past two years was registered.

Multimorbidity

Multimorbidity was computed based on six common conditions across ELSA, HRS and SHARE (arthritis, cancer, chronic lung disease, diabetes, hypertension and stroke). The set of health conditions available from SAGE Wave 0 had only two (arthritis and diabetes) in common; hence the multimorbidity variable was not computed for SAGE.

Social network index

A Social Network Index (SNI) was constructed for ELSA and HRS based on four dimensions: (i) marital status (married/partnered); (ii) religious attendance (at least once a month); (iii) participation in non-religious organizations (at least once a month); and (iv) number of close friends (at least one friend). Each variable was dichotomized and the SNI score could range from 0 (no ties) to 4 (high level of ties).

Subjective well-being: quality of life, life satisfaction and well-being domain

A life satisfaction question was harmonized for ELSA, HRS and SHARE. A question on self-perceived social well-being using a response 'ladder' was available only for ELSA and HRS.

Data resource use

Weights, clustering and stratum variables

Each study provided its own weighting variables, along with the stratum and cluster variables, as part of the complex survey design. This information was used to create a weighted dataset for cross-sectional analysis for SAGE+ Wave 1 (2004).

Statistical analyses

Descriptive statistics were expressed as mean and standard error (SE) for quantitative variables and as age-standardized prevalence for qualitative variables. Weights, strata and cluster information from the SAGE+ harmonized weighted dataset for cross-sectional analysis was used in the *proc surveyreg* or *surveymeans* procedure (SAS v9.2) which produces estimates from complex sample survey data.

Differences in the socio-demographic characteristics

Large samples of adults aged 50+ years were available: SAGE, $n = 18\,886$; ELSA, $n = 9181$; HRS, $n = 19\,303$; and SHARE, $n = 29\,917$. **Table 1** shows national differences in

Table 1. Socio-demographic and economic characteristics: age-standardized prevalence ratios and mean value for SAGE+ Wave 1 (year 2004), for SAGE, ELSA, HRS and SHARE

	SAGE (W0) (n = 18886)	ELSA (W2) (n = 9181)	HRS (W7) (n = 19303)	SHARE (W1) (n = 29917)
Sex				
Male	44.0	47.3	46.5	46.2
Female	56.0	52.7	53.5	53.8
Mean age, years (SE)	62.1 (0.11)	63.9 (0.04)	63.0 (0.07)	63.0 (0.04)
Age groups (years, weighted)				
50–60	41.4	31.1	41.0	36.3
60–70	32.0	31.1	27.8	30.2
70–80	20.8	23.9	19.9	22.3
80+	5.7	14.0	11.3	11.1
Marital status				
Never married	3.1	5.5	4.5	7.8
Currently married/cohabiting	66.0	69.1	64.9	67.2
Separated/divorced/widowed	30.9	25.4	30.5	25.0
Years of education				
0	31.8	0.2	0.5	6.9
1–5	17.6	0.0	2.1	15.5
6–8	19.0	0.0	5.7	19.1
9–13	20.6	75.3	51.5	38.9
14+	10.9	24.5	40.1	19.5
Mean number of years of education (SE)	5.8 (0.16)	11.5 (0.19)	12.7 (0.08)	9.7 (0.14)
Living arrangement				
Alone	13.2	28.7	28.7	31.1
With one person	15.4	50.8	43.0	42.6
With two people	33.3	19.9	26.2	25.3
With three people	23.8	0.6	2.0	1.0
With four people or more	14.3	0.0	0.2	0.1
Mean number of people residing in the household, respondent included (SE)	5.0 (0.11)	2.0 (0.01)	2.2 (0.02)	2.1 (0.02)
Occupation				
Employed	40.3	37.9	45.9	31.8
Unemployed	1.0	1.1	2.8	4.2
Retired	34.0	42.9	32.3	44.5
Homemaker	18.8	10.2	9.0	15.9
Ill health/disabled	4.9	7.0	8.7	3.3
Other	0.9	0.9	1.3	0.3
Mean age at retirement, years (SE)	NA	57.1 (0.14)	58.6 (0.27)	55.2 (0.18)

Prevalence rates are age-standardized using [50–64], [65–74] and [75+] age groups using the WHO standard population. (Ahmad *et al.*, 2001).
NA, not available.

socio-demographic characteristics, with more women in SAGE and ELSA having the highest prevalence of married/cohabiting respondents. Education level varied considerably across studies.

Except for Russia, whose mean number of household members (2.1) is aligned with ELSA, HRS and SHARE, mean household size in the other SAGE countries varies from 3.6 (China) to 7.0 (India).

Considerable variability was found in percentage employed. SHARE reported an unemployment rate as high as 4.2%, HRS and SAGE had similar 'retired' employment status, compared with ELSA and SHARE. Mean age at

retirement was similar across the three studies with available information.

Differences in health status and health habits characteristics

Table 2 reports national differences in health status and habits. Poor-bad-very bad health was reported by 27% in SAGE, compared with 7.5% in ELSA, 7.9% in HRS and 8.1% in SHARE. HRS respondents had the highest prevalence of all chronic health conditions, except for angina, which was highest in SAGE. Hypertension prevalence

Table 2. Health status and health habits characteristics: age-standardized prevalence ratios and mean value for SAGE+ Wave 1 (year 2004), for SAGE, ELSA, HRS and SHARE

	SAGE (W0) (n = 18886)	ELSA (W2) (n = 9181)	HRS (W7) (n = 19303)	SHARE (W1) (n = 29917)
Self-reported health				
Excellent-very good	8.6	41.4	43.3	25.8
Good	23.6	30.5	30.7	40.2
Fair-moderate	39.9	19.6	18.0	25.2
Poor-bad-very bad	27.0	7.5	7.9	8.1
Health conditions (ever been diagnosed/told by a doctor)				
Angina	20.7	9.0	10.2	NA
Arthritis	32.9	34.9	52.3	20.5
Asthma	9.0	12.4	NA	4.1
Cancer	NA	6.8	12.1	5.1
Diabetes	6.9	8.1	16.2	10.3
Hypertension	NA	41.1	49.0	31.4
Lung diseases	NA	6.4	8.9	5.1
Stroke	NA	4.2	4.9	3.2
Measured hypertension	NA	27.3	53.2	NA
Among hypertensive:	NA			NA
Hypertensive not on treatment		15.4	5.8	
Normotensive on treatment		46.4	58.1	
Hypertensive on treatment		38.2	36.1	
Number of chronic conditions (multimorbidity)				
None	52.1 ^a	39.6	23.5	47.8
1	31.5	37.2	32.2	34.6
2+	16.4	23.2	44.3	17.6
Able to remember				
Yes, with no difficulties	38.5	21.8	33.4	NA
With some difficulties	25.9	43.0	41.3	
With difficulty	18.1	28.5	21.0	
With a lots of difficulties	17.5	6.7	4.3	
Time orientation (number of correct items)				
4	NA	80.3	83.7 ^b	86.2
2-3		18.8	15.3	12.6
0-1		0.8	1.0	1.1
Immediate recall (number of words correctly recalled)				
10	NA	0.5	0.7	0.3
7-9		34.6	29.1	16.4
4-6		55.6	61.4	59.5
1-3		8.6	8.1	21.3
None		0.7	0.7	2.5
Delayed recall (number of words recalled)				
10	NA	0.3	0.4	0.2
7-9		14.3	15.0	4.7
4-6		56.2	59.7	40.0
1-3		22.5	20.7	43.6
None		6.6	4.1	11.4
Feeling depressed/sad				
Yes	57.9	26.7	23.8	40.4
Trouble sleeping				
Yes	57.0	42.6	42.3	32.7
Felt lonely				
Yes	NA	13.1	16.1	NA

(continued)

Table 2. Continued

	SAGE (W0) (n = 18886)	ELSA (W2) (n = 9181)	HRS (W7) (n = 19303)	SHARE (W1) (n = 29917)
Able to read				
Yes	51.7	53.5	41.2	37.1
With some difficulties	24.9	35.8	43.2	38.2
With difficulty	11.2	7.9	10.4	15.8
With a lots of difficulties	12.2	2.6	5.1	8.8
Able to seeing things at a distance				
Yes	47.8	60.0	47.8	47.8
With some difficulties	23.9	31.9	41.0	39.4
With difficulty	13.3	5.8	8.1	9.2
With a lots of difficulties	14.9	2.3	3.1	3.6
Hearing	NA			
Excellent		20.7	20.7	17.2
Very good		28.0	27.3	25.6
Good		31.3	33.2	39.1
Fair		15.5	14.6	14.7
Poor		4.5	4.2	3.3
Often troubled with pain				
Yes	73.2	37.4	47.8	50.7
BMI				
Underweight	19.2	0.7	1.2	1.2
Normal	49.3	26.4	31.2	38.9
Overweight	21.2	43.4	39.4	42.6
Obese	10.3	29.4	28.2	17.2
BMI				
mean value (SE)	23.3 (0.11)	28.0 (0.06)	27.9 (0.08)	26.4 (0.07)
Tobacco consumption				
Not current smoker	74.2	82.6	82.9	81.8
Current smoker	25.8	17.4	17.1	18.2
Mean number of cigarettes (SE)	10.4 (0.51)	14.8 (0.28)	15.1 (0.28)	16.1 (0.27)
Number of years with smoking				
Current smoker	37.3 (0.39)	43.0 (0.58)	29.9 (0.71)	NA
Ex-smoker	NA	20.9 (0.23)	19.4 (0.26)	NA
Number of years quit smoking	NA			
Ex-smoker		24.3 (0.36)	21.7 (0.27)	NA
Alcohol consumption				
Abstainer				
Yes	65.5	10.4	44.5	31.1
No	34.5	89.6	55.5	68.9
Heavy drinker				NA
Yes	3.3	1.0	4.3	
No	96.7	99.0	95.7	
Infrequent heavy drinker				NA
Yes	4.5	5.0	3.5	
No	95.5	95.0	96.5	
Number days per week with alcoholic drinks				NA
Less than once a week/once a week	77.1	20.0	54.2	
Twice/three times a week	13.4	35.9	20.6	
Four/five times a week	2.6	18.0	8.8	
Six/seven times a week	6.8	25.9	16.3	
Mean number of drinks per day (SE)	1.57 (0.16)	2.0 (0.03)	2.2 (0.04)	NA
Physical activity				
Frequency of participation in vigorous physical activity				
Less than once a week	60.0	71.0	65.0	50.9

(continued)

Table 2. Continued

	SAGE (W0) (<i>n</i> = 18886)	ELSA (W2) (<i>n</i> = 9181)	HRS (W7) (<i>n</i> = 19303)	SHARE (W1) (<i>n</i> = 29917)
Once a week	4.3	10.1	9.2	13.6
More than once a week	35.7	18.9	25.8	35.4
Frequency of participation in moderate/mild physical activity				
Less than once a week	33.2	7.9	9.4	18.5
Once a week	4.5	7.8	15.8	13.6
More than once a week	62.3	84.3	74.8	67.8
Sedentary				
Yes	8.5	5.4	5.4	11.0
No	91.5	94.6	94.6	89.0

NA, not available.

^aMultimorbidity for SAGE, since it is based on four conditions (angina, arthritis, asthma and diabetes), is not comparable with the other countries (based on six conditions).

^bHRS did not administer the time orientation test to participants aged ≤ 64 .

based on measured blood pressure was lower than self-report in ELSA, and hypertension was controlled better in HRS. A considerable proportion of respondents were hypertensive at measurement even with treatment.

For all three mental health status variables (ability to remember, feelings of depression/sadness and trouble sleeping), SAGE had the highest percentages of poor outcomes.

Near and distance vision showed substantial differences across studies, with SAGE participants reporting the highest percentage of having difficulties in reading and seeing things at a distance.

ELSA, HRS and SHARE had prevalence of overweight, varying from 39.4% (HRS) to 43.4% (ELSA) and obesity from 17.2% (SHARE) to 29.4% (ELSA); SAGE had the highest percentage of underweight respondents.

The prevalence of current smokers was similar across ELSA, HRS and SHARE and was higher than in SAGE, with a higher mean intensity of smoking in these three studies than in SAGE. Drinking habits also varied substantially across studies. Almost 90% in ELSA were not abstainers but heavy drinkers were highest in HRS.

Differences in physical functioning and performance tests

Table 3 shows the prevalence of ADL and IADL limitations, and of specific aspects of mobility/functioning. Any level of dependency (limitation in one or more ADL or IADL) varied between 9.3% in SHARE and 18.6% in ELSA for ADLs, and from 14.3% in SHARE to 23.7% in HRS for IADLs. Grip strength measurements were similar across studies.

Differences in social network and life satisfaction

The social network index (SNI) and life satisfaction variables had many missing data, as indicated in Table 4

(footnote). HRS administered these questions through the self-completion questionnaires with a 20% response rate. About 15% of HRS participants had a high level of ties (SNI = 4) compared with 3.4% in ELSA. The SNI mean value was higher in HRS (2.49, on scale of 0 to 4) than ELSA (1.82), with most satisfied with their life (82.7% in ELSA, 81.7% in HRS and 89.4% in SHARE).

Strengths and weaknesses

Few papers addressed the issue of the ex-post harmonization of ageing studies. Minicuci³ was one of the first attempts to harmonize longitudinal ageing studies from five European countries and Israel. That effort mainly focused on outcomes such as mortality and physical functioning. More recently, the Dynamic Analyses to Optimize Ageing (DYNOPTA) project⁵ constructed a pooled dataset from nine Australian longitudinal studies and focused on outcomes such as dementia and cognition, mental health, sensory impairment and mobility/activity limitations. Crimmins¹⁵ examined gender differences in BMI, smoking status, ADL, IADL, functioning, seven selected diseases and self-rated health among ELSA, HRS and SHARE.

Alternatively, some multi-country studies have been designed with an ex-ante harmonization approach, including the 1999/2000 Survey on Health, Well-Being, and Aging in Latin America and the Caribbean (SABE) examining health conditions and functional limitations of urban-dwelling persons aged 60+ in seven countries,¹⁶ and the COURAGE in Europe study¹⁸ examining the determinants of health and disability in ageing populations in three European countries. Each harmonization effort results in increased data use, thereby improving the return on the considerable financial, human and other resources required for these large studies.^{18–20} They respond to the need for making the best use of secondary data and they allow

Table 3. Physical functioning and performance test characteristics: age-standardized prevalence ratios and mean value for SAGE+ Wave 1 (year 2004), for SAGE, ELSA, HRS and SHARE

	SAGE (W0) (<i>n</i> = 18886)	ELSA (W2) (<i>n</i> = 9181)	HRS (W7) (<i>n</i> = 19303)	SHARE (W1) (<i>n</i> = 29917)
Activities of Daily Living (ADLs) limitations	NA			
Independent		81.4	86.1	90.7
Mild		13.6	10.0	7.0
Moderate		3.7	2.7	1.4
Severe/cannot do		1.3	1.2	0.9
Independent ADLs limitations	NA			
Independent		79.4	76.3	85.7
Mild		15.6	20.0	10.3
Moderate		3.8	2.5	2.3
Severe/cannot do		1.1	1.1	1.7
Mobility				
Walking 100 metres	NA	11.2	10.8	9.1
Sitting for about 2 h	NA	14.4	18.7	10.1
Getting up from a chair after sitting for long periods	NA	25.2	35.5	17.4
Climbing several flights of stairs without resting	NA	35.1	37.5	25.8
Climbing one flight of stairs without resting	NA	13.7	13.4	10.0
Stooping, kneeling, or crouching	NA	35.9	39.4	27.6
Reaching or extending your arms above shoulder level	NA	11.0	13.9	8.1
Pulling or pushing large objects like a living room chair	NA	17.1	19.9	12.7
Lifting or carrying weights over 10 pounds/5 kilos	NA	23.2	17.6	16.8
Picking up a small coin from a table	NA	4.7	5.4	3.0
Walking test	NA			
Mean time over 2.5m seconds (SE)		3.5 (0.03) ^a	3.8 (0.09) ^b	4.9 (0.14) ^c
Grip strength	NA			
Mean (kg) in dominant hand (SE)		30.4 (0.13)	32.7 (0.31)	32.8 (0.22)

All mean values are only weighted and not ag- standardized.

NA, not available.

^aELSA did not administer the walking test to participants aged ≤ 59.

^bHRS did not administer the walking test to participants aged ≤ 64.

^cSHARE did not administer the walking test to participants aged ≤ 74.

Table 4. Social Network Index and life satisfaction characteristics: age-standardized prevalence ratios or mean values for SAGE+ Wave 1 (year 2004), SAGE, ELSA, HRS and SHARE

	SAGE (<i>n</i> = 18886)	ELSA (<i>n</i> = 9181)	HRS (<i>n</i> = 19303)	SHARE (<i>n</i> = 29917)
Social Network Index ^a	NA			NA
0 (no ties)		1.2	0.6	
1		29.9	13.2	
2		58.2	37.0	
3		7.3	34.4	
4 (high level of ties)		3.4	14.8	
Mean SNI		1.82 (0.01)	2.49 (0.03)	
Life satisfaction ^b	NA			
Satisfied		82.7	81.7	89.4
Dissatisfied		17.3	18.3	10.6

NA, not available.

^aELSA has about 40% missing data; HRS has 86% missing data since items on organizational membership and close relationship with friends have been asked in the self-completion questionnaire (response rate around 20%).

^bHRS has 84% missing data since the item has been asked in the self-completion questionnaire (response rate around 20%).

comparable findings across countries to inform public-policy and highlight intervention areas where each country could benefit from another country's policy experiences.

A main strength of harmonized datasets relies on the large sample size achieved and the increased power that allows for more robust and reliable analyses.

Limitations occur when approaching the interpretation of the ex-post harmonization findings. Methodological issues such as generalizability are addressed by the use of specific weights, but for the purpose of cross-countries comparison, the different response rates should also be taken into account. Moreover, differences in self-reported health status are difficult to interpret because these may reflect social and cultural biases and the availability and accessibility of the health services, which distinguish the high-income countries from the low-middle-income countries.

Data resource access

The SAGE+ Wave 1 microdata, along with further details about the harmonization process and all metadata, are available through the WHO data archive at [<http://apps.who.int/healthinfo/systems/surveydata/index.php/catalog>]. Further information and enquiries can be made to [sagesurvey@who.int] or the corresponding author.

Profile in a nutshell

- The generation of a harmonized dataset, comprising a set of variables from four large studies on ageing and health, represents an important step to ensure comparability of common determinants of selected outcomes.
- Although heterogeneous across studies, the common high prevalence of risk factors, such as obesity, tobacco and alcohol consumption and physical inactivity, constitute an important issue for discussion.
- The SAGE+ Wave 1 microdata, along with further details about the harmonization process and all metadata, are available through the WHO data archive.

Supplementary data

Supplementary data are available at *IJE* online.

Funding

SAGE+ was supported through the US National Institute on Aging through grant R21AG034263, 'Harmonizing health outcomes and determinants across longitudinal studies on aging'. SAGE is supported by WHO and the Division of Behavioral and Social Research

(BSR) at the National Institute on Aging (NIA) through Interagency Agreements (OGHA 04034785; YA1323-08-CN-0020; Y1-AG-1005-01) with WHO, and a Research Project Grant R01AG034479. In addition, the governments of China, Ghana and South Africa have provided financial or other support for their national studies.

Competing interests. The authors declare that they have no competing interests.

Acknowledgements

SAGE+ gratefully acknowledges that respondents from each of the studies have given considerable time and energy to this research. The study teams have also given time and expertise to ensure the results are disseminated and data used to inform policy and practice. Mr Yunpeng Huang worked tirelessly with the harmonization team to generate the required metadata documentation.

References

1. Banks J, Marmot M, Oldfield Z, Smith JP. Disease and disadvantage in the United States and in England. *JAMA* 2006;295:2037–45.
2. Suzman R. The INDEPTH WHO-SAGE multicentre study on ageing, health and well-being among people aged 50 years and over in eight countries in Africa and Asia. *Glob Health Action* 2010;3(Suppl 2):5–7.
3. Minicuci N, Noale M, Bardage C *et al.* CLESA Working Group. Cross-national determinants of quality of life from six longitudinal studies on aging: the CLESA project. *Aging Clin Exp Res* 2003;15:187–202.
4. International Household Survey Network. *Metadata Standards – IHSN Technical Note*. Godalming, UK: Metadata Technology, 2012.
5. Anstey KJ, Byles JE, Luszcz MA *et al.* Cohort Profile: The Dynamic Analyses to Optimize Ageing (DYNOPTA) project. *Int J Epidemiol* 2010;39:44–51.
6. Hamilton CM, Strader LC, Pratt JG *et al.* The PhenX toolkit: Get the most from your measures. *Am J Epidemiol* 2011;174:253–60.
7. World Health Organization. *Global Recommendations on Physical Activity for Health*. Geneva: WHO, 2011.
8. World Health Organization. *Global Status Report on Alcohol and Health 2014*. Geneva: WHO, 2014.
9. Kowal P, Somnath C, Nirmala N *et al.* Data Resource Profile: The World Health Organization Study on global AGEing and adult health (SAGE). *Int J Epidemiol* 2012;41:1639–49.
10. Marmot M, Banks JR, Blundell R *et al.* (eds). *Health, Wealth, and Lifestyles of the Older Population in England: The 2002 English Longitudinal Study of Aging. Report No.: P2058*. London: Institute for Fiscal Studies, National Centre for Social Research, 2007.
11. Steptoe A, Breeze E, Banks J, Nazroo J. Cohort Profile: The English Longitudinal Study of Ageing. *Int J Epidemiol* 2013;42:1640–48.
12. National Institute on Aging. *Growing Older in America. The Health and Retirement Study*. Bethesda, MD: NIH Publications, 2007.
13. Sonnega A, Faul JD, Ofstedal MB, Langa KM, Phillips JWR, Weir DR. Cohort Profile: The Health and Retirement Study (HRS). *Int J Epidemiol* 2014;42:576–85.

14. Börsch-Supan A, Brandt M, Hunkler C *et al.* Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). *Int J Epidemiol* 2013;**42**:992–1001.
15. Crimmins EM, Kim JK, Solé-Auró A. Gender differences in health: results from SHARE, ELSA and HRS. *Eur J Public Health* 2011;**21**:81–91.
16. Albala C, Lebrão ML, León Díaz EM *et al.* The Health, Well-Being, and Aging ('SABE') survey: methodology applied and profile of the study population. *Rev Panam Salud Publica* 2005;**17**:307–22.
17. Leonardi M, Chatterji S, Koskinen S *et al.*, on behalf of COURAGE in Europe Project's Consortium. Determinants of Health and Disability in Ageing Population: The COURAGE in Europe Project (Collaborative Research on Ageing in Europe). *Clin Psychol Psychother* 2013;**21**:193–98.
18. National Research Council. *Preparing for an Aging World: The Case for Cross-national Research*. Washington DC: National Academy Press, 2001.
19. National Institute on Aging. *Regional and International Differences in Health and Longevity at Older Ages (R01)*. Bethesda, MD: NIH Publications, 2010.
20. Hense H-W. When size matters. *Int J Epidemiol* 2011;**40**:5–7.