Health and Retirement Study (HRS) Around the World (ATW) Network

Approaching Sample Refreshment after COVID-19

November 20, 2023 • 11:00 a.m. ET

Virtual Meeting

Revised January 4, 2024

This meeting summary was prepared under contract to the University of Michigan (U24 AG037866). The views expressed in this document reflect both individual and collective opinions of the meeting participants and not necessarily those of the University of Michigan. Review of earlier versions of this meeting summary by the following individuals is gratefully acknowledged: Jackie Basu, Ghada Tafesh-Edwards, Rehana Gubin, Rose Maria Li.

Meeting Summary

Welcome and Overview

David Weir, Health and Retirement Study (HRS)

This virtual meeting centers on methods of sample replenishment for longitudinal studies. All longitudinal studies experience diminishing sample size over time due to attrition and mortality. This issue has recently been exacerbated by increased rates of excess mortality due to COVID-19. Representatives of four longitudinal studies—the Mexican Health and Aging Study (MHAS), the English Longitudinal Study of Ageing (ELSA), the Survey of Health, Ageing and Retirement in Europe (SHARE), and the Health and Retirement Study (HRS)—shared sample replenishment experiences and strategies. The list of participants is included as **Appendix A.**

Presentations on Key Topics

Attrition and Mortality in MHAS and the Need for Replenishment

Rebeca Wong, MHAS

MHAS has completed six waves between 2001 and 2021. Since 2012, the study has been conducted at three-year intervals. The sample was replenished for the first time in 2012 and is now replenished at six-year intervals. The seventh MHAS Wave and third sample replenishment will be conducted in 2024. The MHAS team is considering how best to conduct sample replenishment in 2024 to address the impacts of COVID-19 on the sample. To date, sample replenishment has targeted the younger age groups in the study. Each MHAS wave contains approximately 15,000 participants, although the sample size generally increases in sample-refreshed years and decreases in follow-up years.

Effects of COVID-19 on the Sample

The sixth MHAS Wave was conducted in 2021, after the peak of COVID-19 but prior to the Omicron surge. The 2021 sample size is comparable to that of other years when a sample replenishment did not occur. However, the 2021 Wave shows significant increases in next-of-kin interviews across age groups, indicating an increased death rate between the 2018 and 2021 waves, which is consistent with COVID-19-related excess mortality. During the same time, rates of nonresponse, loss to follow-up, and refusal to respond decreased across age groups in the 2021 post-COVID-19 wave.

The weighted MHAS 2021 sample of the Mexican population is closely comparable to 2020 Census data across age groups and genders. However, comparing the *unweighted* MHAS 2021 sample to 2020 Census data and the 2024 National Population Council of Mexico (CONAPO) projections reveals that the MHAS sample underrepresents the population in the 53–69 age groups and overrepresents the 70+ age groups. This comparison signals higher rates of excess mortality in the sample's younger population.

Replenishment of the MHAS Sample in 2024

Both before and after COVID-19, the MHAS sample has consistently overrepresented the older age range. Thus, sample replenishment in 2012 and 2018 focused on the younger population, and the pre-COVID-19 plan for 2024 was to replenish only the younger age groups. Now, the 2024 replenishment will include a new sample of the 50–55 age group. In addition, the MHAS team is considering whether the 2024 replenishment should include older age groups to support mortality studies. However, the team is concerned that (1) allocating sample replenishment to older age groups will weaken the effect of replenishing the underrepresented younger age groups, and (2) a sample replenishment of the older population may have limited effect on longitudinal studies of aging because these participants are unlikely to live long and the study lacks data on their midlife years.

Discussion

Weir noted that there is an option to add to the 55–69 age range and not to the 70+ age range. Sampling weights offer a useful diagnostic to guide replenishment efforts, as populations with larger sampling weights should be prioritized in replenishment. Weir inferred that the younger age ranges were likely to have larger sampling weights and require replenishment, whereas the older population would not. Weir also noted that replenishing the younger age ranges would support longitudinal studies of mortality because the population would contribute data to the study over a relatively long period of time. **Paola Zaninotto** supported Weir's suggested approach to sample replenishment, adding to the 55–69 age range. She also observed that the overrepresentation of the older population could reflect a healthy survivor effect whereby MHAS respondents remain in the study because they are healthier than the general population. However, she concluded that weighting the sample corrects for the slight overrepresentation of the older population.

ELSA's Motivations for Replenishment and Its Experience of Implementation

Paola Zaninotto, ELSA

ELSA has conducted 10 waves and 6 replenishments between 2002 and 2023. Study participants are drawn from the cross-sectional Health Survey for England (HSE). Four of the six replenishments (Waves 3, 6, 7, and 9) targeted the younger age range. Wave 4 replenished the broader 50–74 age range to address both survey dropout and participant mortality. The ELSA team excluded the 75+ age range from Wave 4's replenishment because the higher mortality rate associated with this group would make it difficult to leverage the longitudinal power of the study.

The Wave 10 replenishment also spanned a broad age range; in addition to replenishing the 50–64 age range, it also replenished the ethnic minority groups in the 65+ range to ensure that it accurately represented the non-White British population. The Wave 10 replenishment was planned prior to COVID-19 and scheduled for 2020; however, it was ultimately conducted between 2021 and 2023. Wave 10 response rates were relatively low following COVID-19.
The ELSA sample uses two distinct types of weighting: (1) a *cross-sectional* weight modeled from HSE data, and (2) two sets of *longitudinal* weights, one that began with the Wave 1 cohort

and one that began with the large replenishment sample in Wave 4. As with MHAS, the unweighted ELSA Wave 9 (2018–2019) sample overrepresented the older age ranges and underrepresented the younger age ranges. However, the weighted ELSA Wave 9 sample was closely comparable across ages and genders to mid–2018 English population estimates. These comparisons indicate that the ELSA team's weighting strategies effectively adjusted for bias in the sample. Similar comparisons for Wave 10's sample have not been generated.

Discussion

David Hussey and **Lina Lloyd** noted that the ELSA team has encountered several recent challenges using the HSE. First, HSE data were either unavailable or compromised during the pandemic. Thus, Wave 10's replenishment had to use older HSE data (2017–2018), which led to decreased response rates due to changes in participant contact information. Second, the HSE has become smaller with less available sample, which contributes to an underrepresentation of younger age ranges.

For the upcoming Wave 11, the ELSA team will use the Family Resources Survey (FRS), which offers a larger sample size and greater prevalence of ethnic minorities than the HSE. This will cause the ELSA sample to contain (1) individuals in the same cohort who enter the survey at different times, and (2) individuals in the same cohort who enter via different prior surveys. Weir observed that the HSE and the FRS may draw distinct groups of respondents. The ELSA team has discussed how best to weight this complex sample.

A significant number of individuals included in Wave 10's refreshment either did not respond or could not be covered because of COVID-19 and fieldwork restrictions. Thus, to improve sample response for Wave 11's refreshment, the ELSA team will reissue about 4,000 of Wave 10's HSE refreshment sample and issue 2,000 in the FRS sample.

Lloyd observed that UK age data are not easily accessible. Thus, the ELSA team draws participants from prior surveys to sample from specific age categories. Deriving the ELSA sample from surveys also provides the team access to survey data that participants have consented to share. **Shaun Scholes** observed that not all participants in initial cross-sectional surveys (i.e., either HSE or FRS) agreed to further contact. Thus, the ELSA sample reflects two distinct stages of selective participant responses: (1) consent to further contact following an initial cross-sectional survey, and (2) response to ELSA itself.

The meeting participants discussed country-specific challenges and opportunities in accessing and using survey data. ELSA team members are currently negotiating with UK's Department for Work and Pensions for permission to use FRS data. Weir remarked that federal surveys in the United States prohibit the sharing of participant data for follow-up studies. MHAS may be able to access cross-sectional surveys through its affiliate, the Instituto Nacional de Estadística y Geografía (INEGI).

The ELSA team clarified that they apply longitudinal weights only to individuals who participate continuously in the study. Participants who return to the study after missing one or more waves

do not receive a longitudinal weight, although they do receive a cross-sectional weight. Wong noted that the MHAS team will contact the ELSA team to discuss ELSA's cohort weighting strategy.

Sample Replenishment in SHARE

Michael Bergmann, SHARE

SHARE, which spans 28 European countries and Israel, experienced an interruption during Wave 8 in March 2020 due to the onset of the COVID-19 pandemic. This interruption prompted the inclusion of two additional COVID-19 waves (SCS1 and SCS2) in 2020 and 2021 before regular data collection resumed in Wave 9 in 2022. The study places a significant emphasis on refreshment samples as a crucial mechanism to counteract selective attrition and survivor bias, thereby aiming to enhance the overall sample size.

SHARE countries use different types of sampling frames. In many countries, individuals' information, including age, is obtained from population registers; in the other countries, the individuals' information is lacking, so the team has to screen samples, and oversampling is not always possible due to funding issues.

The study sampling strategies are: (1) a stratified sample of missing youngest cohorts (ages 50-54) only (e.g., a pure refreshment sample), (2) a random sample of the whole 50+ population, and (3) a stratified sample of the 50+ population, including an oversampling of the youngest cohorts. Each of these strategies is tailored to address the unique challenges posed by the diverse demographic and sampling frames across the participating countries.

SHARE's sampling procedure begins with an evaluation of the current net panel sample. Eurostat information for age groups and gender is then used to inform the adaptation of the sample to align with country-specific population figures. Adaptation factors include funding constraints and the types of sample frames or units available. National survey agencies implement the sampling process with a focus on maintaining the study's representativeness across samples.

The study employs the pooling method to address the non-zero cross-selection probabilities between refreshment samples and the existing panel sample. This method adjusts weights to account for individuals' probabilities of being drawn into either the panel or the refreshment sample. Employing distinct weights for individuals and households further refines the precision of analyses. Additionally, applying specialized weights to samples collected during the SCS1 and SCS2 waves acknowledge potential variations in that target population.

Discussion

Weir asked whether the study team compared new recruits and longer-tenured sample members in a longitudinal study, with a specific focus on areas such as cognition, health measures, or wealth. Bergmann noted the lack of extensive research that compares different cohorts, attributing this gap to the common focus on specific cohorts rather than a

comprehensive representation of the entire population. Weir highlighted potential biases in longitudinal samples compared to the general population, noting variations in different settings. He mentioned the challenge of obtaining a representative cross section of older individuals, particularly with conditions such as dementia. Thus, examining factors, such as cognition and health, and comparing new and older recruits of the same age within a study is an important opportunity for a better understanding of the study outcomes. **Jonathan King** noted that work from the Berlin study about 10 years ago concluded that this issue does need to be taken into account.

Replenishment of Minority Cohorts in HRS

David Weir, HRS

HRS has undergone a complex longitudinal cohort design evolution, prominently addressing concerns related to study size and costs. Initially, the study began in 1992 with the original HRS cohort, reinterviewed every two years. Supplemental cohorts, such as the Assets and Health Dynamics of the Oldest Old cohort, were introduced before the adoption of a steady-state design in 1998. This design involved recruiting cohorts aged 51 and older in a given year, waiting six years for aging, and then introducing a new six-year cohort.

The “60 percent solution” was then implemented to recruit new cohorts at 60 percent of the size per birth year of the original HRS cohort. However, this approach led to foreseeable challenges, including future reductions in overall study size, particularly in the minority (Black and Hispanic) samples. The concern about the statistical adequacy of smaller sample sizes prompted a strategic shift. In 2010, HRS secured funding to boost minority sample sizes. This endeavor’s replenishment efforts included recruiting more minority cases for the mid baby boomers and replenishing the early baby boomers (EBB). An analysis of sampling weights in 2004 and 2010 illustrated the impact of these efforts. The EBB cohort, already oversampled in minorities in 2004, saw a significant drop in minority weights in 2010 after introducing additional cases. Replenishment of the EBB cohort resulted in new members with relatively low weights, reflecting the varied selection probabilities in different areas.

The strategy continued with the same sample design in 2016, maintaining an expanded minority oversampling rate in the late baby boomers cohort. In 2022, the study persisted with similar designs for the Early Generation X cohort that undertook additional minority sampling for older cohorts. This ongoing effort will likely necessitate further adjustments in sample weights to accommodate changes in cohort sizes and to maintain the integrity and representativeness of the study, particularly regarding minority groups.

Discussion

Zaninotto asked for clarification regarding the study’s weighting strategies for different cohorts. Weir emphasized the study’s focus on cross-sectional weights for each wave, which avoids the implementation of longitudinal weights. Researchers are advised to start with the weight from the base year relevant to their analysis, modeling attrition and non-response as part of the analytical process. While acknowledging the value of a balanced panel for trends, Weir

suggested that longitudinal weights may not be necessary for most research within the community. He also emphasized the significance of replenishment to improve cross-sectional representation, even though replenishment may not directly impact longitudinal analyses in the short term.

Hiram Beltran-Sanchez (MHAS) highlighted the challenge of accurately representing deceased individuals in mortality studies, in which weights typically reflect the living population interviewed consistently across waves. The difficulty lies in conceptualizing weights that capture the unique dynamics of mortality rates. Predicting the representation of currently alive individuals differs from estimating the representation of those who have died, introducing uncertainties related to potential future deaths. Weir recognized the value of next-of-kin exit interviews to gather data on deaths. He proposed developing and assigning weights to represent deaths in specific HRS age groups or birth cohorts. However, Beltran-Sanchez noted that implementing this approach faces significant challenges in regions such as Mexico, where a reliable age of death registration system is notably absent.

Weir highlighted the importance of internal weighting, particularly to mitigate demographic overrepresentation. He suggested applying internal weighting to assess COVID-19 mortality, which involves comparing the pre-COVID-19 and pandemic-specific weights to establish a weighted mortality rate within specific age groups. This comprehensive approach aims to provide a more refined understanding of mortality patterns, especially in the context of a significant health crisis, such as the COVID-19 pandemic.

Closing Comments

Wong commented that the meeting was useful and thanked participants for their feedback. Weir observed that this meeting helped participants understand specifics about sampling plans beyond reading one another's papers. Zaninotto remarked that the studies often take similar approaches, even when rationales are country-specific or differ based on funders. ELSA technical reports that explain samples and weighting procedures are available on its website (https://www.elsa-project.ac.uk/). **Jinkook Lee** noted that the Longitudinal Ageing Study in India will seek more advice from the group when its refreshment strategy for Wave 3 is developed soon.

Appendix A: Meeting Attendees

U.S. Health and Retirement Study (HRS)
David Weir, PI, HRS; Co-PI, HCAP Network, Co-I, HRS-HCAP; University of Michigan

Chile Cognitive Aging Study (Chile-Cog)
Jere Behrman, Co-PI, University of Pennsylvania

China Health and Retirement Longitudinal Study (CHARLS)
Yafeng Wang, Peking University

English Longitudinal Study of Ageing (ELSA)
Paola Zaninotto, Co-PI and Deputy Director; Co-Investigator, Gateway to Global Aging Data
David Hussey, Team Member, National Centre for Social Research
Lina Lloyd, ELSA Management Group, National Centre for Social Research (NatCen)

Irish Longitudinal Study on Ageing (TILDA)
Ann Hever, Research & Development Manager

Longitudinal Aging Study in India-Diagnostic Assessment of Dementia (LASI-DAD)
Jinkook Lee, PI, LASI-DAD; University of Southern California

Mexican Cognitive Aging Ancillary Study (Mex-Cog)
Rebeca Wong, PI, Mex-Cog; University of Texas Medical Branch (UTMB)
Hiram Beltran-Sanchez, UCLA
Alejandra Michaels-Obregon, Lead Research Project Manager, UTMB

Panel Survey on Health, Aging, and Retirement in Thailand (HART)
Dararatt Anantanasuwong, PI
Pailin Chuayok, Team Member

Survey of Health, Ageing and Retirement in Europe (SHARE)
Michael Bergmann, Head of Survey Methodology
David Richter, Director SHARE Infrastructure
Sime Smolic, Croatia Country Team Leader

Understanding America Study (UAS)
Arie Kapteyn, PI

National Institute on Aging (NIA), National Institutes of Health (NIH)
Minki Chatterji, HCAP Network Project Officer; Program Official, Population and Social Processes Branch, Division of Behavioral and Social Research (DBSR)
Jonathan King, HCAP Network Project Scientist; HRS and HRS Around the World Project Scientist; Senior Scientific Advisor to the Division Director, DBSR
Shanna Breil, Social Science Analyst, DBSR

Rose Li and Associates, Inc. (RLA)
Sofia Jones, Meeting Planner