The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 2025, **80**(S1), gbae150
https://doi.org/10.1093/geronb/gbae150
Advance access publication: 31 August 2024
Supplement Article: Studying Cognition in the National Social Life, Health, and Aging Project: Methodological and Analytic Considerations, Strategies and Findings From Round 4

Transitioning an In-Person Survey of Older Adults to Multimode Data Collection

Anna Wiencrot, MPH,[1,*] Katie O'Doherty, MA,[1] Daniel Lawrence, MA,[1] Sara Walsh, MA,[1] Jennifer Satorius, MSW,[1] Lauren Sedlak, MPP,[1] Kelly Pudelek, MA,[1] Kaitlyn O'Keefe, BS,[1] Evelyn Zepeda, BS,[1] Hanvit Park, MPP,[1] Joscelyn Hoffmann, BA,[2] Hannah You, BA,[2] Martha McClintock, PhD,[3] and Stephen Smith, MA[1]

[1]NORC at the University of Chicago, Chicago, Illinois, USA.
[2]The Institute for Mind and Biology, University of Chicago, Chicago, Illinois, USA.
[3]Department of Psychology, University of Chicago, Chicago, Illinois, USA.

*Address correspondence to: Anna Wiencrot, MPH. E-mail: Wiencrot-anna@norc.org

Decision Editor: Hui Liu, PhD (Social Sciences Section)

Abstract

Objectives: The first 3 rounds of the National Social Life, Health, and Aging Project (NSHAP) were in-person. Preparing for Round Four (R4), NSHAP began developing ways to collect complex questionnaire and biomeasure data remotely. R4 was scheduled to begin in 2020, but due to the coronavirus pandemic, NSHAP delayed R4 data collection and instead conducted a study on respondents' experiences during the pandemic, as well as pretests to strengthen NSHAP's remote data collection capability. This paper describes the methodology, results, and lessons learned from these efforts, which were undertaken as a bridge between NSHAP's all-in-person past and multimode future.

Methods: The coronavirus disease 2019 (COVID-19) Study was a multimode survey of NSHAP respondents to assess the impact of the pandemic. The multimode approach allowed evaluation of the feasibility of using different modes of data collection with older adults. NSHAP adapted its in-person questionnaire for phone and web administration and conducted pretests of the full phone questionnaire and sections of the web questionnaire. The project developed and tested a "BioBox," a kit containing all the supplies and instructions for respondents to self-collect biomeasures remotely. The BioBox was tested through an in-lab and in-home pilot, followed by 2 larger-scale pretests.

Results: The COVID-19 Study and pretests achieved NSHAP respondent participation in remote questionnaire and biomeasure collection, despite being accustomed to fully in-person data collection.

Discussion: Our findings and experiences will inform the collection of NSHAP data in future rounds and could inform other panel studies of older adults considering multimode data collection.

Keywords: Aging, Biomeasures, Health, NSHAP, Survey

Overview of Data Collection Waves

The National Social Life, Health, and Aging Project (NSHAP) is a longitudinal, population-based study of health and social factors that provides information about the well-being of older adults by examining interactions among physical health, illness and disease, cognitive function, emotional and mental health, and social connectedness. The NSHAP interview consists of three components: (1) a detailed in-person questionnaire focusing on physical, mental, and social health; (2) the collection of minimally invasive biomeasures; and (3) a leave-behind paper questionnaire (LBQ). The in-person interviews occur in the respondent's home and are administered via Computer-Assisted Personal Interview. NSHAP also collects biological measures, prioritizing noninvasive collection techniques and innovative technologies that minimize respondent burden (see Jaszczak et al., 2009). The NSHAP sample is a national area probability sample of community-residing older adults and consists of two cohorts: (1) Cohort 1 includes adults born between 1920 and 1947 and their coresident romantic partners of any age; and (2) Cohort 2 includes adults born between 1948 and 1965 and their coresident romantic partners of any age.

In 2005–2006, Round 1 (R1) of NSHAP was conducted, recruiting Cohort 1 from a nationally representative sample of adults in the United States aged 57–85 at the time of recruitment. R1 was designed to: (a) describe the health of older Americans; (b) evaluate the relationship between older adults' health and social connections; (c) examine the importance of gender and socioeconomic context for older adults' health; (d) identify biological pathways through which social connectedness affects aspects of health; and (e) describe health behaviors and healthcare utilization among older adults. NORC at the University of Chicago interviewers completed 3,005 interviews in R1, achieving a weighted response rate of 75.5% (see Smith et al., 2009).

To expand the initial data, to allow for longitudinal analyses of long-term health trajectories, and to better characterize

Received: March 28 2024; Editorial Decision Date: August 19 2024.

the impact of relationships on health, interviewers returned in 2010–2011 for Round 2 (R2), interviewing Cohort 1 respondents from Round 1 (then aged 62–90) and, for the first time, interviewing their cohabiting spouses and romantic partners. Along with the in-person interview and LBQ, the NSHAP R2 interview added, removed, and enhanced certain biomeasures (see O'Doherty et al., 2014). Interviewers completed 3,377 interviews in R2, achieving a weighted conditional response rate of 89% for R1 respondents and 84% for their partners, as well as a 26% conversion rate for people who were invited to participate but declined in R1 (see Jaszczak et al., 2014; O'Muircheartaigh et al., 2014).

Round 3 of NSHAP (R3) focused on continuing the longitudinal analyses of long-term health trajectories and outcomes by reinterviewing Cohort 1 respondents (then aged 67–95), as well as recruiting Cohort 2 respondents (aged 50–67) for their first interview to begin analyzing differences across generations. R3 was conducted in 2015–2016 and included Cohort 1 respondents and their coresident romantic partners (n = 2,409) and a cohort of new respondents and their coresident romantic partners (n = 2,368). NORC developed a short household screening questionnaire to determine the eligibility of one or more members of the sampled household to recruit this new, younger sample of Cohort 2 respondents. R3 included the household screener, the in-person interview, and the collection of biomeasures. Interviewers completed 4,777 interviews in R3 (see O'Doherty et al., 2021) for a 91% conditional response rate for Cohort 1 and a response rate of 76% for Cohort 2 (see O'Muircheartaigh et al., 2021). Figure 1 summarizes both cohorts across the first three rounds of NSHAP.

Overview of the Pretests and Studies Bridging R3–R4

The fourth round of NSHAP data collection was intended to be conducted in 2020–2021, continuing to examine long-term health trajectories and outcomes by reinterviewing both cohorts of NSHAP respondents. Because multimode surveys were becoming increasingly common, the National Institute on Aging requested that for R4, NSHAP develop a version of the interview that could be conducted remotely. In 2019, NSHAP conducted the first two of these remote adaptations in the form of a Phone Questionnaire Pretest and the collection of self-administered biomeasures using a kit called the "BioBox" in the First BioBox Pretest. Building on the success of these efforts, NSHAP began preparing for R4, which was intended to be mostly in-person interviews with a subset of respondents participating in the new remote version.

Rather than implementing these 2019 adaptations immediately in R4 2020 as planned, R4 data collection was delayed by the coronavirus disease 2019 (COVID-19) pandemic. During this delay, we fielded an interim COVID-19 Study with NSHAP respondents. This short survey asked about health and relationships during the pandemic and was available by web, phone, and paper. Simultaneously, we continued adapting the remote protocol to be administered more broadly than originally planned. This work consisted of a Second BioBox Pretest and a Web Questionnaire Pretest. At the end of the COVID-19 Study questionnaire, some respondents were invited to the Second BioBox Pretest while a different set of respondents were invited to the Web Questionnaire Pretest. Figure 2 presents a timeline and brief description of the pretests and COVID-19 Study. The goal of this paper is to describe the methodology, results, and lessons learned of the work the NSHAP team undertook between the conclusion of R3 and the beginning of R4 to adapt, test, and refine NSHAP's extensive in-person data collection protocols for remote, multimode administration.

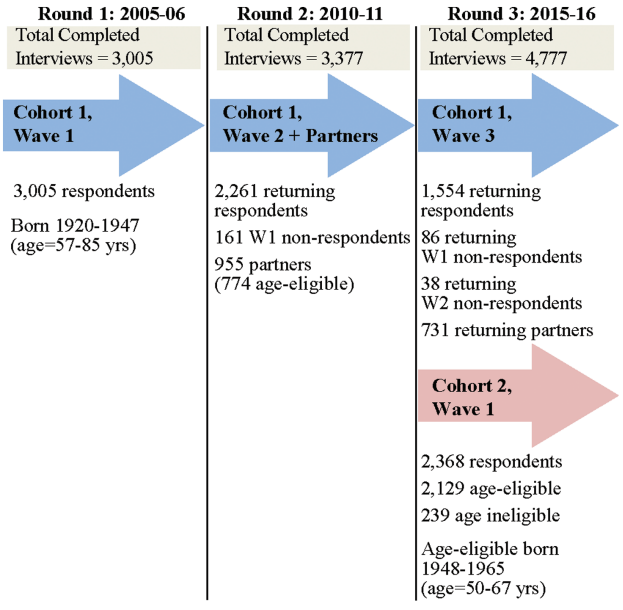

Figure 1. NSHAP respondent participation Rounds 1–3. NSHAP = National Social Life, Health, and Aging Project.

COVID-19 Study

Beginning in the fall of 2020, NSHAP fielded a COVID-19 Study to study the impact of the COVID-19 pandemic on NSHAP respondents. Respondents answered questions about their health and relationships during the pandemic, administered by web, phone, or paper. The project also used this study as an opportunity to make the remote protocol more robust to be administered to more respondents than originally planned for R4.

COVID-19 Study Sample and Design

The COVID-19 Study sample comprised 4,852 respondents (2,531 Cohort 1 and 2,321 Cohort 2), all of whom had participated in at least one previous round of NSHAP. Data collection was conducted from September 2020 through January 2021. All eligible respondents were first mailed an invitation to participate in the web questionnaire. Those who did not respond to the web invitation after 4 weeks were sent a paper questionnaire. Nonrespondents to web and paper after were called on the phone starting in Week 10. A final paper questionnaire was sent to all nonrespondents in Week 14. Inbound phone was available throughout the field period. Modes were presented sequentially, although respondents could still complete the survey in preceding modes after new ones were opened. For example, some respondents chose to complete the survey by web after receiving the paper questionnaire by mail.

COVID-19 Study Questionnaire

The COVID-19 Study questionnaire included some questions typically asked in NSHAP as well as new questions developed

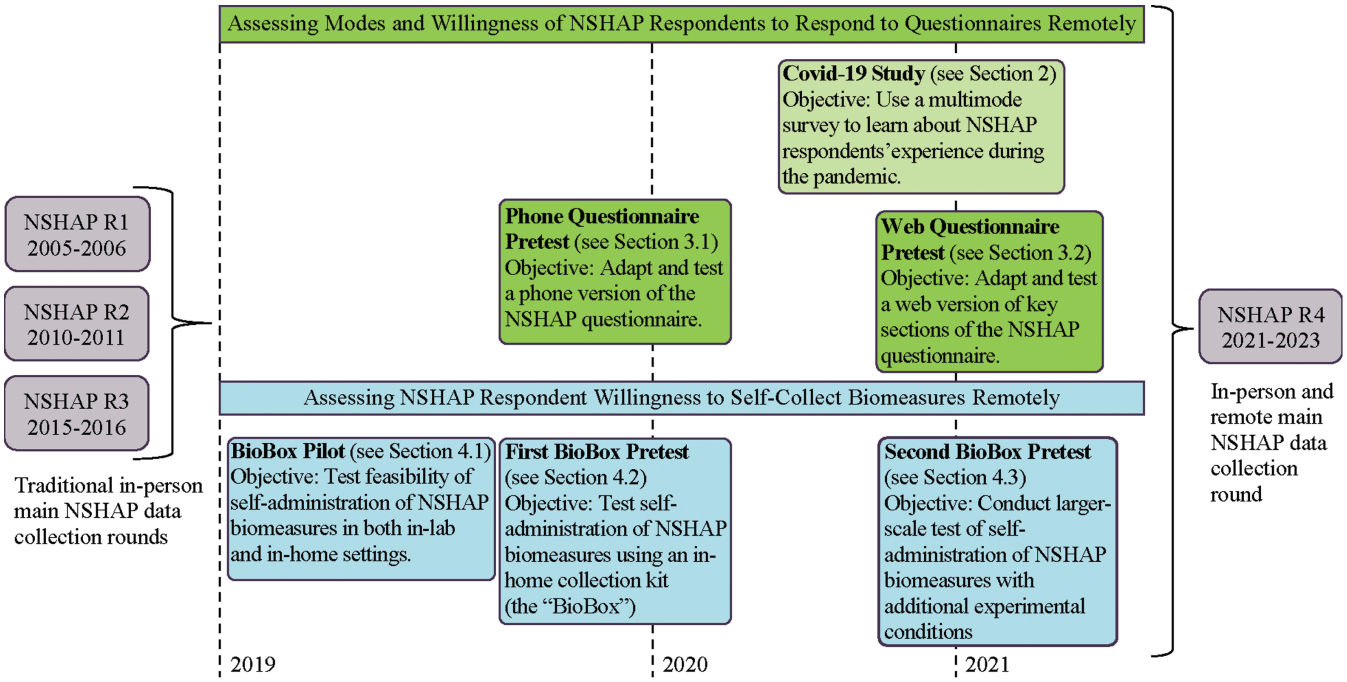

Figure 2. Timeline of adapting NSHAP data collection from in-person to multimode. NSHAP = National Social Life, Health, and Aging Project.

to assess how respondents' lives had changed or stayed the same since the start of the pandemic. The questionnaire contained approximately 60 questions and took approximately 20 min to complete. Each question was assessed for compatibility across all three modes and the wording of questions or response options was adjusted as needed to accommodate mode delivery. For example, self-administered modes (web and paper) used first-person (e.g., "my spouse") while the phone questionnaire used second-person (e.g., "your spouse"). Optional prompts and clarifications were also added to the phone questionnaire for interviewers to use as needed. The structure of some questions also differed by mode, such as adding leading language for phone questions that were presented as grids in web and paper. Web and phone respondents were shown consent language within the survey and asked to confirm they would like to continue while the paper questionnaire provided consent language on the cover. At the end of the instrument, some web respondents were asked to participate in a separate Web Questionnaire Pretest of the regular NSHAP instrument (see *NSHAP Web Questionnaire Pretest*). Another group of respondents who completed the survey by *any mode* (web, phone, or paper) was asked to participate in the Second BioBox Pretest (see *Assessing NSHAP Respondent Willingness to Self-Collect Biomeasures Remotely*).

COVID-19 Study Gaining Cooperation

NSHAP used a multipronged approach to gain respondent cooperation. Outreach was conducted in multiple modes (e.g., hardcopy letter, e-mail, phone) to ensure awareness and underscore the importance of the study. Multimode data collection encouraged and facilitated participation in the respondent's preferred mode. NORC interviewers were trained to answer common questions and received guidance on how to respond to respondent answers and comments. In response to lower response rates to the COVID-19 Study, Spanish-speaking respondents were sent a tailored outreach letter that included responses to FAQs addressing confidentiality and privacy concerns.

COVID-19 Study Results

Table 1 shows the final status of all NSHAP COVID-19 Study cases by cohort. A total of 2,656 respondents, or 55% of the sample, completed the questionnaire. Both cohorts responded at about the same rate, with 55% of Cohort 1 completing the survey and 54% of Cohort 2. Among the noninterviewed respondents, only 4% explicitly refused to participate, while most (36%) we were simply unable to contact.

Table 2 shows the NSHAP COVID-19 Study completes and partial completes by mode. Just over half (51%) of respondents participated by web. Paper, the second mode offered to respondents, made up 39% of the completes. The remaining completes were conducted over the phone, with approximately half of the phone interviews resulting from inbound calls and half from outbound.

NSHAP's younger respondents (Cohort 2) responded at a higher rate to web (59%) than Cohort 1 (44%), while older respondents (Cohort 1) completed by phone (14%) and paper (43%) at higher rates than Cohort 2 (8% by phone and 34% by paper).

Assessing Modes and Willingness of NSHAP Respondents to Respond to Questionnaires Remotely

Prior to the pandemic, NSHAP had begun developing the capacity for remote data collection. In preparing for R4, which would be the first round with remote data collection, the team adapted the typical in-person NSHAP questionnaire for phone administration, with web and paper being added after the success of those modes in the COVID-19 Study.

Table 1. Final COVID-19 Study Case Status by Cohort

	Cohort 1		Cohort 2		All	
Status	*n*	%	*n*	%	*n*	%
Complete	1,395	55%	1,261	54%	2,656	55%
Partial complete	8	0.3%	8	0.3%	16	0.3%
Breakoff	6	0.2%	21	0.9%	27	0.6%
NIR—refused	128	5%	86	3.7%	214	4%
NIR—all others (e.g., no contact)	852	34%	912	39%	1,764	36%
Out of scope	142	6%	33	1%	175	4%
Total	2,531	100%	2,321	100%	4,852	100%

Notes: COVID-19 = coronavirus disease 2019; NIR = noninterviewed respondent.

Table 2. Final COVID-19 Study Completes by Mode

	Completes		Partial completes		All	
Mode	*n*	%	*n*	%	*n*	%
Web	1,340	51%	14	88%	1,354	51%
Inbound phone	145	6%	0	0.0%	145	5%
Outbound phone	139	5%	2	13%	141	5%
Paper	1,032	39%	0	0.0%	1,032	39%
All	2,656	100%	16	100%	2,672	100%

Note: COVID-19 = coronavirus disease 2019.

Phone Questionnaire Pretest

Adapting the NSHAP questionnaire for phone

The first step in developing the Phone Questionnaire Pretest was to review the in-person NSHAP questionnaire and ensure all items could be collected accurately and easily over the phone by interviewers, while minimizing respondent burden. Steps to ensure this included: (1) removing cognitive measures that had a visual or interactive component; (2) creating an interactive script for interviewers to help respondents enumerate their social network roster (SNR); and (3) carefully reworking response options where in-person interviewers would typically rely on show cards. Most sections and questions were able to remain the same as in-person or only needed minor modifications.

The NSHAP interview includes tasks that measure areas of cognitive functioning such as attention, memory, visuospatial orientation, and language. Some tasks have a visual component, which is not possible on the phone. These were removed to create a shorter cognitive section that could be completed on the phone. All other cognitive items remained the same.

The SNR section of the questionnaire asks respondents to enumerate people with whom they discuss important things, assesses respondent marital/partnership status, assesses frequency of contact between network members, and asks respondents to compare their list of people with a list of people they named in previous interviews. During the typical, in-person NSHAP interview, an interviewer guides the respondent with a physical roster page and show cards. For the phone, we created a script that asked the respondent to get a piece of paper and write down their roster as they named the members. Some questions in the SNR require show cards due to the considerable number of response options, such as when asking the respondent's relationship with roster members. For these questions, we created a shorter list of summary response options, with optional follow-up probes for the interviewer to identify the correct response.

Table 3. Final Phone Questionnaire Pretest Dispositions

Final distribution of Phone Questionnaire Pretest dispositions	All respondents	Total %
Phone completes	48	39.0%
Final noninterviewed respondent (NIR)—refusal	12	9.8%
Final NIR—no contact/unlocatable	61	49.6%
Final out of scope (deceased, too poor health)	2	1.6%
Total	123	100%

Notes: NIR = noninterviewed respondent.

Phone Questionnaire Pretest study design and sample

These adaptations for phone were tested during the Phone Questionnaire Pretest in late 2019. Because the remote protocol was originally intended for Cohort 2, the sample selected was approximately the age of Cohort 2 respondents. The 123 pretest respondents were drawn from a list sample licensed from Marketing Systems Group; some had participated in previous NSHAP pretests and some were new to NSHAP. The data collection period was 11 weeks, from November 2019 to January 2020, and was conducted by three trained interviewers.

Phone Questionnaire Pretest gaining cooperation

Interviewers reported that although it was often difficult to reach respondents by phone, once they spoke to a respondent, they were usually able to avert or convert refusals by explaining the study's importance and offering to send additional study materials. Interviewers found that texting was sometimes effective for reaching respondents who could not be reached by phone.

Phone Questionnaire Pretest results

Table 3 shows the final dispositions of the 123 cases approached. A total of 48 respondents, or 39%, completed the questionnaire. The median phone interview time was 64 min, and the mean phone interview time was 69 min. The phone questionnaire functioned as intended and interviewers reported no significant challenges beyond occasional difficulty with them hearing the respondent or the respondent hearing them. This hearing issue was most apparent when the respondent was asked during the cognition section to say words starting with a certain letter.

NSHAP Web Questionnaire Pretest

The next phase in the process of adapting NSHAP data collection for remote administration was to develop and test a self-administered web version of the questionnaire. It was critical to conduct a larger-scale feasibility test of this mode because the NSHAP sample is older adults, some of whom we hypothesized might be less technologically proficient or

simply less inclined to participate by web. In addition, some questionnaire items required significant adaptation to be administered without an interviewer.

Adapting the NSHAP questionnaire for web

As with the phone questionnaire, we reviewed the in-person questionnaire to ensure all measures could be answered by web accurately and easily, and with minimal burden. Most sections required very few changes. However, three major questionnaire sections posed unique challenges for web: (1) the SNR, (2) cognition, and (3) a medication log.

As discussed above, the SNR section asks respondents to enumerate people with whom they discuss important things, assesses respondent marital/partnership status, assesses frequency of contact between network members, and asks respondents to compare their current list of people with a list of people they named in previous interviews. During the typical in-person NSHAP interview, an interviewer uses a physical roster sheet to elicit names from the respondent, so we added detailed instructions to help the respondent understand how to fill in the names and follow-up questions without the assistance of an interviewer.

The NSHAP interview also includes tasks that assess areas of cognitive functioning such as attention, memory, visuospatial orientation, and language (see Pudelek et al. for additional information on measuring cognition in R4, this issue). The challenge in adapting this section for the web is that certain measures could only be done on paper. We also needed to ensure that respondents could only view each question or stimulus once to maintain the comparability of these measures to in-person data (e.g., the interviewer reads—and cannot repeat—a list of five words and asks the respondent to remember and repeat them); to achieve this, the "back" button was disabled for certain questions in this section. In addition to adapting existing measures, we incorporated another web-based cognition tool (ManyBrains) to determine its usability within the NSHAP sample. This usability test involved two ManyBrains exercises: one that displayed a series of unrelated word pairs and then asked the respondent to remember them, and one that displayed a series of unrelated photos and then asked the respondent to remember them.

The third component of the NSHAP interview that we tested was the medication log. During the typical in-person interview, the interviewer asks the respondent to gather all their medications, then enters the information from the medication packaging. In moving this to the web, we were concerned that respondents taking many medications might underreport their medications or break off from the questionnaire entirely. In addition to allowing respondents the more traditional option of typing in their medications, we tested a brand-new adaptation that allowed respondents to write their medications on a piece of paper and upload a photo of the paper. We also provided a visual aid for how to locate the information needed on the medication package.

The goal of the Web Questionnaire Pretest was to test these three self-administered web sections with NSHAP's older adult population. After a brief introduction (during which the respondent confirmed their identity) and consent, the interview proceeded in three parts: the respondent answered the substantive questions for each of the three major sections, and after each section received 4–5 questions that asked for feedback on the section they had just completed (e.g., "Was there anything unclear or confusing in this section?").

Web Questionnaire Pretest sample and study design

The sample comprised 450 respondents from the main NSHAP sample who had completed the NSHAP COVID-19 Study by web from September through November 2020. An invitation to the pretest was mailed to eligible respondents that provided the link to the survey. E-mail and mail reminders were subsequently sent to nonresponders and people who started but did not finish the questionnaire. Pretest data collection lasted 7 weeks, from November 2020 to early January 2021.

Web Questionnaire Pretest Results

A total of 334 respondents, or 75% of the sample, completed the questionnaire. Table 4 shows final case status by age range. Most respondents (72%) completed the questionnaire using a laptop or desktop computer (see Author Note 1). Of the 450 respondents invited to participate in the web pretest, 28 respondents started the questionnaire but did not finish it. We were particularly interested in *where* these respondents broke off, as it could indicate confusion or frustration with certain sections. The highest number of breakoffs (22 of 28) occurred during or immediately following the ManyBrains section of the cognition measures. This could potentially be a result of the respondent's experience of ManyBrains, which some respondents described as difficult or frustrating. The ManyBrains section took a significant amount of time in the overall questionnaire, and it required paying close attention to a succession of images and words that many respondents found stressful. Breakoffs also occurred during the medication log. We believe that some respondents took so much time away from their computer to gather their medications that the questionnaire timed out and they either chose not to, or were unable to, log back in.

Another notable finding was that respondents provided fewer people in their SNR compared to NSHAP's R3 in-person

Table 4. Final Web Questionnaire Pretest Status by Age

Final case status	<51		51–60		61–70		71–80		81–90		>90		Total	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Complete	6	86%	43	69%	81	70%	155	79%	44	72%	5	56%	334	74%
Partial complete	0	0%	4	6%	8	7%	12	6%	4	7%	0	0%	28	6%
NIR	1	14%	15	24%	24	21%	29	15%	11	18%	4	44%	84	19%
Out of scope	0	0%	0	0%	2	2%	0	0%	2	3%	0	0%	4	1%
Total	7	100%	62	100%	115	100%	196	100%	61	100%	9	100%	450	100%

Note: NIR = noninterviewed respondent.

data collection (i.e., an average of 3.27 roster members in the Web Questionnaire Pretest vs 4.46 in R3). This was likely due to the way the question was asked (e.g., requesting names one-by-one) rather than providing space to add all names at once. Because of the importance of the SNR data, we were concerned that the lower average number of members added in this pretest could be an artificial result of the web interface rather than a true representation of significant change in respondents' social networks. Based on this finding, it was clear we would need to reconsider how to revise this section when developing the NSHAP R4 study (including potentially moving the collection of all roster members to a single screen).

Web Questionnaire Pretest respondent feedback

A key objective for this pretest was understanding how NSHAP respondents used and experienced the questionnaire. Feedback on the survey was collected formally through questions embedded in the questionnaire, as well as informally when respondents contacted the project by phone or e-mail.

At the end of each section, respondents were asked open-ended questions about their experience with that section. Over 80% of respondents did not report finding anything unclear or confusing with any of the sections. Of respondents who reported issues with at least one section, the majority of those responses were in reference to the cognition section.

- SNR: Feedback received on the SNR section largely concerned comments about the roster itself and the network change section, rather than the SNR's web usability. Although these comments suggest some respondents had difficulty navigating the roster in self-administered mode, most of these comments were linked to the content of the roster, suggesting that in typical NSHAP in-person interviewing, interviewers help resolve such issues.
- Cognitive Measures: Apart from the ManyBrains section, respondents generally found the cognition section to be clear. Feedback for this section focused largely on its usability. When asked whether anything was confusing or unclear about the cognition section, most of the respondent feedback referenced the ManyBrains section. Respondent feedback about ManyBrains made it clear that it did not test well for wider use with the full NSHAP sample in R4.
- Medication Log: Some respondents requested additional fields for entering medications while others found it unpleasant to type in all their medications. Despite this feedback, fewer than 2% of respondents chose the option to upload a picture of their medications list.

Additionally, 65 respondents contacted the project by phone or e-mail during the pretest. The most common issue was difficulty accessing the questionnaire using the URL and/or the PIN and password. We suspect respondents had difficulty typing in the URL, PIN, or password correctly from the invitation letter. These issues were resolved by e-mailing respondents a personalized link with their PIN and password embedded.

Assessing NSHAP Respondent Willingness to Self-Collect Biomeasures Remotely

NSHAP created the BioBox, a mailed kit containing biomeasure supplies and devices, to facilitate respondent self-collection of biomeasures. A specialized print booklet called the BioBooklet provided illustrations and detailed instructions for collecting the measures, plus space to record the measure outcomes. With limited precedent for older adults self-collecting physical measures and biospecimens with such a kit, multiple rounds of development and testing were required. This section describes the initial pilot studies of the BioBox, followed by the First BioBox Pretest and a larger Second BioBox Pretest.

Initial Development and Pilot Study of BioBox

The BioBox was initially developed and tested via a pilot study designed and executed by the NSHAP investigators and supported by the NORC operations team. The BioBox Pilot consisted of both an in-lab phase and an in-home phase. The in-lab phase (January to April 2019) recruited approximately 70 community-dwelling participants aged 50–80 to complete NSHAP biomeasures using protocols adapted for self-administration at the Institute for Mind and Biology (IMB) at the University of Chicago while being observed by NSHAP investigators and IMB staff. The NORC operations team observed in-lab visits to provide operational and data collection feedback. At the end of the in-lab visit, participants provided feedback on the instructions, individual measures, and overall self-collection experience.

The measures included in the in-lab phase of the pilot were Cognition, Weight, Waist, Taste, Touch, Distance Vision, Smell, Hearing, Timed Walk, Chair Stands, Saliva, and Vaginal Swabs using two different devices (see Table 6 for all health measures collected during the phases of BioBox testing). The observational data and results of the self-collected and investigator-administered measures were analyzed by the investigators and used to refine collection protocols and design the prototypes of the BioBox and the BioBooklet (see Author Note 2).

For the second, in-home phase (September to October 2019), prototypes of the BioBox and BioBooklet were sent to respondents. The in-home phase of the pilot study was a small convenience sample of approximately 10 members of the NSHAP Investigator team and their family who were within the target age range (50–80). Feedback from participants and the data were analyzed and used to create the final design of the BioBox and BioBooklet for the First BioBox Pretest.

Outcomes of the pilot included clarification and simplification of protocol instructions for reading comprehension, a color-coding schema for the measure types (used in the BioBooklet and supply labeling), selection of the vaginal swab collection method, exclusion of blood pressure, and adaptation of weight to a self-collected measure. Lastly, to reduce respondent burden, measures were divided into two groups, so each group needed to complete only a subset of the full range of measures.

First BioBox Pretest

First BioBox Pretest sample

The First BioBox Pretest was conducted with the same respondents as the Phone Questionnaire Pretest (see *Phone Questionnaire Pretest*). At the end of the phone questionnaire, respondents were provided information about the BioBox pretest and those who consented were sent the BioBox. All 48 respondents who completed the Phone Questionnaire Pretest consented to receive the BioBox.

First BioBox Pretest BioBox design

The final design of the BioBox for this pretest included two separate groupings of measures (referred to as Schedules) to gauge respondent burden. All respondents were asked to complete sensory measures (Taste, Touch, Vision, and Hearing) as well as body measures (Waist and self-reported Weight). Schedule A respondents were also asked to complete Cognition and the activity measures (Timed Walk, Chair Stands, and leave-behind Accelerometry), while Schedule B respondents were also asked to collect biosamples (Saliva and, for women, Vaginal Swabs; Table 6 displays the measures in each Schedule). In both Schedules, the BioBox package was a white rectangular mailer box with a handle on it, resembling a briefcase, with a University of Chicago label on the outside. The BioBox contained the necessary supplies for completing the measures and returning the equipment, samples, and BioBooklet.

First BioBox Pretest logistics

Respondents who consented received their assigned BioBox version via FedEx. Respondents could either drop their completed BioBox at a FedEx location, call FedEx for a pickup, or have their field interviewer call and arrange a pickup. All BioBoxes were shipped back to NORC's office downtown Chicago and specimens were couriered to the investigator lab daily.

First BioBox Pretest respondent support

Interviewers made three types of calls to BioBox pretest respondents:

1. *Arrival*: This call confirmed delivery of the BioBox, reviewed the BioBooklet, and answered any preliminary questions.
2. *Check-in*: This call touched base with respondents about BioBox progress and tried to get an estimate from them about completion date.
3. *Prompting*: This call followed up with respondents on the date they expected to ship back the BioBox.

In general, after the phone questionnaire portion of the interview was complete, most respondents wanted to avoid phone calls with the interviewer or keep these calls short. Respondents told interviewers that the instructions in the BioBooklet were clear and concise. Moreover, when interviewers called to confirm the arrival of the BioBox and talk through the measures and supplies, most respondents did not feel the need to talk through the details; however, the calls were beneficial for encouraging respondents to open their BioBox and get started. Very few respondents reached out to interviewers with questions or concerns about the measures.

First BioBox Pretest results

First BioBox Pretest cooperation and return rates

All 48 respondents returned their BioBox with at least one measure attempted. The individual cooperation rates of each measure are shown in Table 5. Although all 48 respondents returned their accelerometer after wearing it for 8 days, we learned that respondents struggled to activate the devices and that the devices required additional charge to last the full 8 days. Overall, this pretest of the BioBox showed that respondents were willing and able to collect their own biomeasures. However, the NSHAP Pretest sample (from which these respondents were drawn) has historically been more cooperative than the main NSHAP sample and we did not anticipate the same rates for the main study.

First BioBox Pretest respondent feedback

We contacted a subset of respondents after they had completed the BioBox to validate their experience and collect feedback. Almost all respondents who completed this short debriefing offered positive feedback on the overall experience, their interviewer, project materials, and BioBox. This post data collection conversation yielded the following: (1) respondents reported an average BioBox completion time of 46 min, which is not significantly different than what was reported in the returned BioBooklets; (2) respondents reported that the BioBox instructions were clear and easy to follow; (3) three-quarters of respondents reached said that the amount

Table 5. Biomeasure Cooperation Rates From First and Second BioBox Pretests

Measure	First BioBox Pretest: cooperation rate	Second BioBox Pretest: cooperation rate	In-person R3: cooperation rate
Cognition/Puzzles	100%	99%	99%
Smell	100%	98%	96%
Taste	100%	96%	n/a[b]
Vision	98%	97%	n/a[b]
Touch	96%	95%	n/a[b]
Hearing	n/a[b]	89%	n/a[b]
Timed Walk[a]	91%	90%	98%
Chair Stands[a]	87%	95%	98%
Weight[a]	94%	100%	98%
Waist Circumference	100%	100%	97%
Saliva	84%	90%	91%
Vaginal Swabs	77%	77%	n/a[b]
Accelerometry[a]	100%	72%	84%
Blood[a]	n/a[b]	93%	92%

Notes: NSHAP = National Social Life, Health, and Aging Project.
[a]There were slight differences between this protocol in the BioBox Pretest(s) and in NSHAP R3.
[b]Measure not collected in that pretest/round.

Table 6. Biomeasures Included in the BioBox

	BioBox Pilot	First BioBox Pretest: Schedule A	First BioBox Pretest: Schedule B	Second BioBox Pretest: Schedule A	Second BioBox Pretest: Schedule B	Second BioBox Pretest: Schedule C
Cognition	●	●		●		●
Self-reported Blood Pressure and Pulse				●	●	●
Weight	●					
Self-reported Weight		●	●	●	●	●
Waist	●	●	●		●	●
Taste	●	●	●	●	●	●
Touch	●	●	●	●	●	●
Distance Vision	●	●	●	●	●	●
Smell	●	●	●	●	●	●
Hearing	●	●	●			●
Timed Walk	●	●		●		●
Chair Stands	●	●		●		●
Activity and Sleep/6-min walk		●		●		●
Saliva	●		●		●	●
Vaginal Swab	●		●		●	●
Blood Spots (Tasso)				○	○	○

Notes: ●= respondents were asked to complete this measure; ○ = half of the respondents were asked to complete this measure.

of time they spent on the BioBox was "just right"; (4) about half of respondents reported doing the BioBox measures all in one sitting, while about a third spread the measures out (the remaining respondents did not answer); and (5) no respondents reported difficulty with packing up their BioBox for return shipping or getting the box shipped back.

Second BioBox Pretest

The BioBox was initially developed and designed for respondents in NSHAP's younger (Cohort 2) age range. However, due to the COVID-19 Pandemic, NSHAP sought to: (1) increase the percentage of respondents invited to remote data collection, (2) expand BioBox eligibility to include Cohort 1 respondents, and (3) add additional health measures to reduce the gaps in biomeasure data between in-person and remote data collection modes. Because the BioBox would be offered to more respondents than initially planned, it was important to further evaluate its design with a larger sample, as well as gauge the impact of length and type of measures on cooperation rates in the NSHAP main sample. We conducted the Second BioBox Pretest with respondents who participated in NSHAP's COVID-19 Study, which would provide insight into BioBox cooperation with the main NSHAP sample, rather than the more cooperative than average sample in the First BioBox Pretest.

Second BioBox Pretest sample

A subset of 600 respondents from the NSHAP COVID-19 Study were selected to participate in the Second BioBox Pretest. This subset was separate from the subset invited to participate in the Web Questionnaire Pretest, but like that pretest, all selected respondents were part of the main NSHAP sample. Respondents were randomly assigned one of six BioBox conditions and asked to consent to the BioBox at the end of their COVID-19 Study questionnaire.

Second BioBox Pretest BioBox design

NSHAP investigators designed six BioBox configurations to test which combination of measures would yield the best cooperation rates. The six configurations aimed to assess whether adding additional measures and increasing the overall length of the BioBox would have an impact on cooperation rates, and whether the addition of a self-collected blood measure would have an impact on cooperation rates. The six configurations came from three different Schedules: (1) Schedule A: short BioBox (1 hr) collecting activity measures, (2), Schedule B: short BioBox (1 hr) collecting biospecimens, and (3) Schedule C: long BioBox (1.5 hr) collecting both activity and biospecimen measures. Each of the three Schedules was split into two conditions—one with blood collection and one without blood collection—to form six configurations.

Second BioBox Pretest BioBox contents

All BioBox versions included: Blood Pressure and Pulse (self-reported), Weight (self-reported), Taste, Touch, Distance Vision, and Smell. About half the respondents in each Schedule were additionally asked to collect a blood sample using a novel device called the Tasso. Respondents placed the Tasso on their arm, pressed a button to prick the skin, then a small amount of blood gathered into four pods internal to the device.

In addition to those measures, Schedule A also included Chair Stands, Timed Walk, Cognition, and Activity and Sleep. The Activity and Sleep Measure involved wearing a wrist-worn accelerometer for 8 days and completing a 6-min walk during that time. Schedule B's additional measures included Saliva (via passive drool), Vaginal Swabs (female respondents only), and Waist Circumference. Schedule C included all measures from Schedules A and B. Table 6 presents the measures for the BioBox Pilot and the First and Second BioBox Pretests.

Table 7. BioBox Consent, Return, and Yield Rates in the Second BioBox Pretest

BioBox types	Eligible	Consented	Consent rate	Returned	Return rate	Yield rate
Box A + blood	99	73	73.74%	69	94.50%	69.70%
Box A—no blood	100	75	75.00%	74	98.70%	74.00%
Box B + blood	101	72	71.29%	70	97.20%	69.30%
Box B—no blood	91	62	68.13%	57	91.90%	62.64%
Box C + blood	105	60	57.14%	49	81.70%	46.67%
Box C—no blood	99	73	73.74%	66	90.40%	66.67%
Total	595	415	69.75%	385	92.77%	64.70%

Second BioBox Pretest logistics

Respondents were asked to indicate a time frame that would work best to receive their BioBox. As in the First BioBox Pretest, respondents could either drop their completed BioBox at a FedEx location, call FedEx for a pickup, or have their interviewer arrange a pickup. BioBoxes were shipped to NORC's office downtown Chicago and specimens were couriered to the investigator lab daily.

Second BioBox Pretest gaining cooperation

Respondents received an advance letter with a study brochure that briefly mentioned the BioBox. Once the respondent completed their COVID-19 Study questionnaire, cooperation for the BioBox was gained one of four ways: (1) at the end of the phone questionnaire with a telephone interviewer reading a script, (2) at the end of the web questionnaire with a BioBox information screen, (3) included as a separate sheet with the PAPI mailing, or (4) collected by an interviewer calling respondents who initially declined to receive the BioBox. Descriptions on web, phone, and paper all described the BioBox and its contents. For respondents selected into a condition with a blood collection, they were also informed of the opportunity to receive results from a COVID-19 antibody test to see if antibodies to SARS-CoV-2, the virus that causes COVID-19, were present in their blood sample. In all modes, if a respondent indicated interest in receiving a BioBox they were given a full copy of the consent to read (on web or paper) or have read to them. Telephone interviewers were provided a list of FAQs to help answer questions, and field interviewers had been trained on in-depth knowledge of the BioBox.

Second BioBox Pretest respondent support

When respondents were sent a BioBox, interviewers began outreach to provide support. This outreach included: shipping notifications; checking in to confirm receipt of the BioBox, answer questions, and confirm planned completion date; and following up on missing BioBoxes and accelerometers.

Second BioBox Pretest results

Second BioBox Pretest consent, return, and cooperation rates

Of the 595 eligible respondents, 415 respondents (69.8%) consented to receive the BioBox. A total of 385 respondents, or 92.8% of those who consented, returned their BioBox. Table 7 shows the final status of returned BioBoxes. Of the sample eligible for accelerometry, 64.4% returned the accelerometer after wearing it for 8 days. The Short BioBoxes (Schedules A and B) had higher return rates than the Long BioBox (Schedule C), and the inclusion of blood appeared to depress return rates, particularly in Schedule C with blood.

Table 5 shows cooperation rates by measure for respondents who returned the BioBox. These rates were comparable to those of respondents who completed the in-person NSHAP R3 interview, demonstrating that respondents who returned a BioBox were likely to attempt all measures. It should be noted that the cooperation rates of the in-person respondents are a percentage of all completed interviews, whereas the cooperation rates of the Second BioBox Pretest are a percentage of all respondents who returned a BioBox. The cooperation rates do not capture respondents who completed the interview but declined the BioBox or who completed the interview but failed to return the BioBox.

Second BioBox Pretest respondent feedback

Respondents were asked for feedback on what was fun and what was hard. Of the 372 respondents who provided feedback on what was fun, 120 answered that almost everything was fun. The most common positive themes demonstrated that respondents: enjoyed learning about their current lifestyle and health; appreciated the opportunity to conduct certain measures with their family members; reported the instructions were clear and easy to follow; and were happy to contribute to the study. Respondents reported difficulty with Blood, Fitness Trackers, and Vaginal Swabs. In addition, some respondents reported that packing up their BioBox for return shipping was hard, which was opposite of reports from the First BioBox Pretest.

Discussion

Findings From Questionnaire Testing

In the COVID-19 Study, a significant portion of respondents participated on web, indicating that web was viable for a subset of NSHAP's older adult respondents. However, in both the COVID-19 Study and the Web Questionnaire Pretest, we learned that significant technical support may be needed. Initial outreach via a mailed invitation inviting respondents to complete the web survey required respondents to navigate to the survey website and enter a PIN and password to access their survey. Some respondents struggled with this step. E-mail outreach facilitated participation by offering a direct survey link. However, we had e-mail addresses for less than a third of respondents, and e-mails were easy to ignore (particularly as some were caught in spam filters). An additional limitation of the COVID-19 study is that it occurred during the pandemic—a time of widespread disruption—and therefore respondents may have had more time to complete a survey remotely when so many aspects of commerce and society were paused.

Table 8. Summary of the Studies and Pretests Undertaken by NSHAP to Transition From In-Person to Multimode

NSHAP Pretest/ study	Objective	Time frame	Sample size	Sample frame	Mode(s)	Completion rate
NSHAP COVID-19 Study	Use a multimode survey to learn about NSHAP respondents' experience during the pandemic	September 2020 to January 2021	4,852	Respondents who have participated in at least one round of NSHAP	Web/paper/phone	55%
Phone Questionnaire Pretest	Adapt and test a phone version of the NSHAP questionnaire	November 2019 to January 2020	123	List sample licensed from Marketing Systems Group of age-eligible adults located in Illinois and Indiana	Phone	39%
Web Questionnaire Pretest	Adapt and test a web version of key sections of the NSHAP questionnaire	November 2020 to January 2021	450	A subset of respondents who completed the COVID-19 Study were assigned to either the Web Questionnaire Pretest or Second BioBox Pretest	Web	75%
BioBox Pilot (in-lab phase)	Test feasibility of self-administration of NSHAP biomeasures in an in-lab setting	January 2019 to April 2019	~70	Community-dwelling adults age 50–80 in the Chicago area	Biomeasure self-administration with in-person observation	100%
BioBox Pilot (in-home phase)	Test feasibility of self-administration of NSHAP biomeasures in in-home settings	September 2019 to October 2019	~10	Members of the NSHAP Investigator team and their family age 50–80	Biomeasure self-administration	100%
First BioBox Pretest	Test self-administration of NSHAP biomeasures using an in-home collection kit (the "BioBox")	November 2019 to January 2020	48	List sample licensed from Marketing Systems Group of age-eligible adults located in Illinois and Indiana	Biomeasure self-administration	100%
Second BioBox Pretest	Conduct larger-scale test of self-administration of NSHAP biomeasures with additional experimental conditions	November 2020 to January 2021	595	A subset of respondents who completed the COVID-19 Study were assigned to either the Web Questionnaire Pretest or Second BioBox Pretest	Biomeasure self-administration	65%

Notes: COVID-19 = coronavirus disease 2019; NSHAP = National Social Life, Health, and Aging Project.

The pretests of the Phone and Web questionnaires provided additional opportunities to evaluate how remote questionnaire administration methods differed from in-person administration. Our findings were encouraging about the prospect of remote data collection with some NSHAP respondents. On the phone, lessons learned include providing guardrails against mishearing words over the phone (such as the memory words in the cognition section) and providing more support and guidance to interviewers as they collect consent to certain sensitive measures (such as Medicare number) remotely instead of relying on the personal rapport that an in-person experience cultivates. On the web, significant changes were needed for the SNR, cognition, and medication log. Challenges with cognition on the web underscored the importance of an easy-to-navigate instrument, as well as identifying ways to reduce burden to prevent respondent frustration and breakoffs. In particular, we learned that, for a number of respondents, the ManyBrains tool was long and taxing, and the fact that it opened in a different tab interface than the NSHAP questionnaire caused confusion for some respondents.

Findings From BioBox Testing

The BioBox Pilot and First and Second BioBox Pretests allowed NSHAP to explore the feasibility and best methods for remote biomeasure collection. The cooperation rates achieved indicate a promising future for between-wave visits and remote collection of self-administered biomeasure collection. A future paper will discuss the usability of the measures collected in the BioBox.

The BioBox conditions with fewer measures had the highest response rates, and the inclusion of blood collection reduced the response rates. These were important findings to carry forward into the crafting of the BioBox for NSHAP R4, particularly regarding the need to balance the scientific importance of blood collection against the risk that asking respondents to self-collect blood could reduce cooperation rates.

While most respondents did not request interviewer support to complete the BioBox, the phone prompting protocol answered respondent questions and was effective in prompting the return of the BioBox. Any future scripts for this type of remote gaining cooperation and prompting should allow the interviewer to put the text into their own words to facilitate building trust and communication with the respondent, while also recognizing that certain respondents did not seem to want any additional contact by phone.

Overall, remote biomeasure collection proved feasible with older adults; it was the total number of measures included (rather than age or some other factor) that had the highest impact on overall return rate of BioBoxes, and cooperation

rates were comparable to NSHAP's in-person biomeasure cooperation rates. While the yield rate for remote biomeasure collection is lower than for in-person, respondents who did return the BioBox completed most of the measures. This could indicate that additional resources should be allocated to investigating ways to increase consent to the BioBox.

Closing

To recap, all the data collection episodes described in this paper are part of the overall work NSHAP has undertaken to transition its traditionally in-person-only study to a robust multimode study (see Author Note 3). The most important finding, therefore, was the willingness of NSHAP respondents to participate in a remote interview. Each of these interim efforts to adapt NSHAP for remote administration provided important lessons to be used in the development and fielding of NSHAP R4 and which could serve as a guide to other longitudinal studies seeking to transition into new modes of data collection.

Author Notes

1. Device type collected by questionnaire paradata rather than self-report.

2. A future paper will provide additional information on the BioBox and BioBooklet development process.

3. Table 8 presents a summary of the studies and pretests conducted by the NSHAP team and described in this paper.

Funding

This supplement is supported by the National Institute on Aging (NIA) at theNational Institutes of Health (NIH) (R01AG043538, R01AG048511) and NORC at the University of Chicago, which was responsible for the data collection. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIA/NIH or NORC.

Conflict of Interest

None.

Data Availability

NSHAP data and documentation are available to registered users through the National Archive of Computerized Data on Aging, within the Interuniversity Consortium for Political and Social Research. More details on how to gain access to NSHAP data can be found at https://www.icpsr.umich.edu/web/pages/NACDA/nshap.html. The NSHAP questionnaires and other data collection documentation are publicly available at www.norc.org/nshap.

Acknowledgments

NSHAP is directed by Linda J. Waite, Principal Investigator. Coinvestigators: Edward Laumann, Martha McClintock, Colm O'Muircheartaigh, Kathleen A. Cagney, William Dale, Louise Hawkley, Elbert Huang, Diane Lauderdale, L. Philip Schumm, Benjamin Cornwell, Megan Huisingh-Scheetz, Jayant Pinto, Lissette Piedra, Ashwin Kotwal, and Ellis Monk.

References

Jaszczak, A., Lundeen, K., & Smith, S. (2009). Using nonmedically trained interviewers to collect biomeasures in a national in-home survey. *Field Methods*, 21, 26–48. https://doi.org/10.1177/1525822X08323988

Jaszczak, A., O'Doherty, K., Colicchia, M., Satorius, J., McPhillips, J., Czaplewski, M., Imhof, L., & Smith, S. (2014). Continuity and innovation in the data collection protocols of the second wave of the National Social Life, Health, and Aging Project. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 69(Suppl. 2), S4–S14. https://doi.org/10.1093/geronb/gbu031

O'Doherty, K., Jaszczak, A., Hoffmann, J. N., You, H. M., Kern, D. W., Pagel, K., McPhillips, J., Schumm, L. P., Dale, W., Huang, E., & McClintock, M. K. (2014). Survey field methods for expanded biospecimen and biomeasure collection in NSHAP wave 2. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 69(8), S27–S37. https://doi.org/10.1093/geronb/gbu045

O'Doherty, K., Lawrence, D., Wiencrot, A., Walsh, S., Satorius, J., Burgess, E., Sedlak, L., Koepp, K., & Smith, S. (2021). Ongoing refinement and innovation in the data collection protocols of the third round of the National Social Life, Health, and Aging Project. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 76(Issue Supplement_3), S215–S225. https://doi.org/10.1093/geronb/gbab179

O'Muircheartaigh, C., English, N., Pedlow, S., & Kwok, P. K. (2014). Sample design, sample augmentation, and estimation for wave 2 of the NSHAP. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 69(Issue Supplement_2), S15–S26. https://doi.org/10.1093/geronb/gbu053

O'Muircheartaigh, C., English, N., Pedlow, S., & Schumm, L. P. (2021). Sample design and estimation in the National Social Life, Health, and Aging Project: Round 3 (2015–2016). *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 76(Issue Supplement_3), S207–S214. https://doi.org/10.1093/geronb/gbab182

Smith, S., Jaszczak, A., Graber, J., Lundeen, K., Leitsch, S., Wargo, E., & O'Muircheartaigh, C. (2009). Instrument development and implementation of the study design for the National Social Life, Health, and Aging Project (NSHAP). Special Issue in *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 64(Suppl. 1), 20–29. https://doi.org/10.1093/geronb/gbn013