How Different Mixed-Mode Data Collection Approaches Impact Response Rates and Provision of Biomeasure Samples

Tarek Al Baghal[1],[*] [ID], Jonathan Burton[2] [ID], Thomas F. Crossley[3] [ID], Michaela Benzeval[4] [ID], Meena Kumari[5] [ID]

[1]Professor of Survey Methodology, Institute for Social and Economic Research, University of Essex, Colchester, UK
[2]Associate Director (Surveys) of Understanding Society, Institute for Social and Economic Research, University of Essex, Colchester, UK
[3]Research Professor and Co-Director of Panel Study of Income Dynamics, Institute for Social Research, University of Michigan, Ann Arbor, MI, US
[4]Professor and Director of Understanding Society, Institute for Social and Economic Research, University of Essex, Colchester, UK
[5]Professor and Director, Institute for Social and Economic Research, University of Essex, Colchester, UK

Abstract   Social surveys can be enriched with the collection of objective health measures, allowing new types of research in both health and social sciences. We experimentally tested three alternative designs for collecting survey responses and biomeasures within a longitudinal survey. In the nurse-administered design, a nurse conducts the survey and collects biomeasures in person. In the interviewer-first design, an interviewer initially attempts to carry out the survey in person, collects a subset of biomeasures, and then leaves a further biomeasure sample collection kit with the respondent. The web-first design invites respondents to complete the survey in web mode, and a biomeasure sample collection kit is sent after they do so. Nonrespondents to their initial mode are followed up with in an alternate mode. The outcomes of interest are both (i) response to the survey, and (ii) take-up and completion of the biomeasure sample collection. The impact of the experimental design is tested on both outcomes, utilizing intention-to-treat analysis (that is, by allocated design). To account for the importance of channel of communication in the consent decision for biomeasures, we also analyze observed consent outcomes by realized mode of response, other survey factors, and respondent characteristics. Findings show that the web-first design is superior in obtaining survey response, with nonsignificant differences between in-person interviewer-administered and nurse-administered designs. Conversely, the web was the least effective

*Corresponding author: Tarek Al Baghal, Institute for Social Research, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK; email: talbag@essex.ac.uk.

design for obtaining biomeasures. These findings imply that there is a design trade-off between obtaining survey responses and biomeasures, and this should be considered in future studies.

Introduction

Social surveys can be enriched by the collection of objective health measures (biomeasures). A biosocial survey collects richer health measures than a traditional social survey, and more social and economic measures than a health study. This combination facilitates the study of the social determinates of health, and the effects of health on other life outcomes, such as employment and retirement (Benzeval et al. 2016). Collecting socioeconomic measures and objective health measures from the same respondents can promote the use of survey data in fields such as health economics, epidemiology, behavioral science, health psychology, and medical sociology, and promotes interdisciplinary research (McFall et al. 2012).

Given the benefits of combining objective health measures and a social survey, an important question is how best to conduct such a biosocial survey. The two components are typically collected in quite different ways. For a health study, the preferred design is the use of clinics (Golding 2004; Marmot and Brunner 2005; UK Biobank 2007; Kenny et al. 2010; Kuh et al. 2011; Zipf et al. 2013). This design collects the most comprehensive and high-quality set of biomeasures. Clinics may be centralized or mobile but involve in-person collection by trained medical staff. In contrast, most social surveys are conducted by trained interviewers, in person or by phone, or, more recently, by self-administration on the web. We experimentally test three alternative designs for collecting survey responses and biomeasures within a longitudinal biosocial survey.

Participation in a clinic-based study requires a significant time commitment by the respondent, including travel to the clinic. There is a risk that participation is selective, with those in poorer health being less likely to take part (Kearney et al. 2011). This is particularly undesirable in a large biosocial survey that aims to support population inferences. In addition, clinic-based data collection is expensive. As a result, a number of alternative designs have been developed to collect biomeasures from a general population sample alongside a social survey. These designs avoid bringing respondents to a clinic.

In one such alternative design, trained medical professionals (usually nurses) visit respondents in their own homes. An interviewer first visits the home and completes the social interview with the respondent, and afterward a nurse visits to collect the biomeasures. This sequential design, in which

nurses follow interviewers in separate visits, has been used in the English Longitudinal Study of Ageing (ELSA—Marmot and Steptoe 2008; Steptoe et al. 2013), the Health Survey for England (HSE—Mindell et al. 2012), and Understanding Society (McFall et al. 2012). The range and quality of the biomeasures collected in-home are lower than in clinic-based studies. Nurses, or phlebotomists, may take blood samples, but these often then must be posted, and so may take some days to reach a laboratory. This delay may affect the range of analytes that can be measured robustly (Peakman and Elliott 2008). Importantly, this design reduces the burden on the respondent; the collection is done in their home rather than them having to travel to a clinic. However, the gap in time between the interviewer visit and the nurse visit increases the chance of attrition between the two stages (Clemens et al. 2012). Respondents to the social interview may refuse the follow-up nurse visit, or the nurse may not contact them or persuade them to take part when contact is made. In addition, this design brings the additional cost of two separate visits to the home.

Instead of nurses collecting biomeasures during a subsequent visit, interviewers can carry out the interview and collect a limited set of biomeasures during a single visit (Weir 2008; Jaszczak et al. 2009; Sonnega et al. 2014; Guyer et al. 2017; Weiss et al. 2019). A single visit reduces costs. However, the time in the home for that visit will be longer. This increases the burden for respondents. Moreover, although the interviewers can be trained to collect a range of biomeasures, they are not clinically trained and the types of equipment they can use are limited. For example, they can be trained to collect dried blood spots or saliva samples, but they cannot collect whole blood, which limits the range of possible analyses. A variation on this design would have interviewers instead (or additionally) provide kits for respondents to self-collect biomeasures after the interview, while additionally explaining the purpose and procedures to motivate and aid that subsequent self-collection. Such a design would have the desirable effect of reducing the time the interviewer spends in the home. However, not all respondents will return self-collected samples, which reduces return rates and may increase selection bias.

An alternative possibility is that nurses conduct the social survey. This requires a single visit, and so is less costly than the sequential interviewer-with-nurse-follow-up design, and eliminates attrition between visits. It also allows for a wider range of biomeasures to be collected by a nurse, while still collecting the social survey data. However, the increased number of biomeasures requires a longer interview, and so may be more of a burden for the respondent. Nurses may be less effective in obtaining survey responses than trained interviewers. This design was used in the first half of the 2016 sweep of the 1970 British Cohort Study (Brown et al. 2019). In that instance, nurses had lower survey response rates than interviewers, and the nurse

design was discontinued mid-wave, given the lower survey response rate and higher costs.

Finally, social surveys are often conducted by telephone or by web, without any home visit. This approach can be extended to a biosocial survey by asking respondents to self-collect the biomeasures after completing the telephone interview or web survey, and to return the biomeasure sample. Obvious concerns with this design include the ability and willingness of respondents to self-collect the biomeasures (without in-person instruction); and whether respondents will return the samples. While mixed-mode designs (using both self-completion and interviewer modes) can obtain higher survey response rates than single-mode designs (De Leeuw 2018; Olson et al. 2021), obtaining biological samples depends directly on the channel of communication (Dykema et al. 2016). Interviewers can use tailoring and persuasion techniques to obtain consent, and are able to answer questions from respondents, which is not possible in a self-administration design. Table 1 summarizes these alternative designs for biomeasure collection.

Dykema et al. (2016) used a telephone survey with a saliva collection kit sent in a follow-up mailing and obtained a 54 percent kit return. Chartier et al. (2021) also used a telephone survey to introduce a saliva collection kit, and obtained a 53 percent response and consent rate, with 60 percent of those who consented to receive the saliva kit returning it. Note that in these two cases the survey was interviewer administered, but by telephone. Web surveys are entirely or primarily self-administered. The collection of blood and saliva samples was piloted on the Dutch LISS internet panel in 2010, with around 15 percent of respondents returning samples (Avendano et al. 2010). Gatney et al. (2013) achieved a return rate of 65 percent for saliva samples from the mixed-mode (web with telephone follow-ups) Relationship Dynamics and Social Life (RDSL) study. There has been a recent increase in the use of self-collection of biomeasures, largely driven by the COVID-19 pandemic. During the pandemic, web participants in Understanding Society: UKHLS were asked for their consent to be sent a blood sample kit to test for COVID-19 antibodies, with just under 80 percent giving consent (Burton et al. 2024). Several other large studies measured the prevalence of COVID-19 in the UK via respondent-led biomeasure collection (Riley et al. 2021; Ghafari et al. 2024).

While the literature reports the implementation of these designs, there are few studies that directly compare the effectiveness of different designs. Jaszczak et al. (2009) found that requests for respondent self-collection of biomeasures performed worse than interviewer collection. Consent rates are lower in web modes than interviewer-administered modes (Sakshaug et al. 2017; Jäckle et al. 2021). On the other hand, Allen et al. (2024) found that there was a *higher* proportion of samples collected from an at-home group (74.4 percent) than in a clinic setting (58.4 percent), although both had

Table 1. Alternative designs for biomeasure collection in surveys.

	Advantages	Disadvantages
Clinic based	Wide range of physical, cognitive, biological measures possible. High-quality measures by trained clinicians.	High burden for respondents—long biomeasure collection time, interview time, and need to travel to clinic. High cost.
Nurse visit	Nurses specialize in biomeasure collection. Just one visit to respondent required, so reduces cost and burden.	Narrower range of biomeasures than in a clinic. Collection in a nonclinical setting, so lower quality than clinic based. Nurses may be less successful in obtaining survey response.
Interviewer with nurse follow-up	Interviewers specialize at getting social survey interview and obtaining acceptance of nurse follow-up. Nurses specialize in biomeasure collection.	Narrower range of biomeasures than in a clinic. Collection in a nonclinical setting, so lower quality than clinic based. Risk of attrition between interviewer and nurse visit. Respondent is required to agree to two visits to their home. High cost of visits by interviewer and nurse.
Interviewer collection	Just one visit to respondent required, so reduces cost and burden compared to two visits.	Interviewers can collect a narrower range than nurses. Collection in a nonclinical setting so lower quality. Interviewers have some training, but not as much as nurses. Interview time will be longer, may lead to respondent fatigue.
Interviewer provide kit for self-collection	One visit to respondent required, shorter than combined social and biomeasure collection, so reduces burden. Interviewers can explain the process to respondents more thoroughly and answer questions.	Narrower range of biomeasure possible, lower quality of measures (untrained respondent). Respondents may not return samples.
Respondent collection	Respondents can collect biomeasures when convenient to them. Cheap to collect.	Narrower range of biomeasure possible, lower quality of measures (untrained respondent). Respondents may not return samples.

similar return rates (93.3 percent and 94.5 percent, respectively). In terms of quality, Fuller et al. (2019) found that in-home blood sample collection by participants resulted in a higher proportion of good quality samples than those taken by staff in clinics. However, these studies were not implemented experimentally, and so it is not possible to isolate the effect of the mode and other design features on response and return of biomeasures.

We contribute to filling this gap in the literature by experimentally testing three alternative designs for collecting survey responses and biomeasures within the Innovation Panel of Understanding Society: UKHLS. The designs we test vary by whether they involve an in-person home visit, and by how (or by whom) the interview and biomeasure collection are administered. The three designs we test are:

  i. Nurse led
  ii. Interviewer first
  iii. Web first

The designs we test in this paper are sequential mixed-mode designs, in which nonrespondents in one mode may be followed up in another. Using a sequential mixed-mode design may achieve improved survey outcomes (De Leeuw 2018; Rybak 2023; Sakshaug et al. 2019). Employing multiple modes facilitates diverse initial contact channels, such as combining advance letters with emails. Such strategies are designed to improve response rates (De Leeuw 2005).

In this paper, we address the following research questions:

RQ1: Do the different designs differ in the response rate achieved to the annual social survey?
RQ2: How do the different designs impact the return of biological samples?

Theoretical Framework

Survey outcomes are influenced by myriad factors, with survey mode playing a crucial role in shaping these outcomes (Bowling 2005). Mode will also be crucial in both the survey response rate and the provision of biomeasures. Mode of survey administration can affect factors such as respondent engagement, trust, and burden, and these in turn impact respondents' willingness to take part in the survey and/or return biomeasures. Understanding these dynamics is essential for optimizing desired outcomes. Designs that foster engagement and trust and that reduce burden will likely improve both response rates and biomeasure provision. However, survey mode can influence these factors in different and possibly opposite ways in relation to survey response and biomarker provision. In this context, we discuss how mode affects these factors, as well as how

logistical practicalities should be considered. By examining these jointly, we aim to show how different survey modes can strategically enhance response rates and biomeasure provision, thereby improving data quality and reliability.

Engagement

Interviewers, through their interpersonal skills, can effectively engage respondents, leading to higher response rates for the survey itself and possibly for biomeasure provision (Sakshaug et al. 2010; Korbmacher and Schroeder 2013). Interviewers can build rapport, tailor their approach to the respondent, and use persuasive techniques to encourage participation. In contrast, nurses, who may not possess the same level of training in these areas, might struggle to achieve similar engagement levels, potentially affecting overall response rates (Brown et al. 2019). While web surveys can incorporate some level of tailoring, this is generally limited compared to face-to-face interactions (Bianchi et al. 2017). Additionally, the absence of immediate feedback and clarification may result in lower engagement and higher dropout rates.

Trust

Respondents' trust can be fostered or hampered by the survey mode, which can significantly impact response rates and biomeasure provision. Interviewer and nurse-led surveys can more effectively address privacy concerns through personal interaction, thereby fostering trust and possibly improving response rates. Nurses' medical expertise can enhance trust and willingness to provide biomeasures due to their perceived authority and ability to address specific health-related questions or concerns (Kearney et al. 2011). Interviewers lack medical expertise, however, which may limit engendering trust and understanding of the biomeasure component of the study. The impersonal nature of web interactions may make it challenging to alleviate concerns about survey requests (De Leeuw 2005), including biomeasure provision.

Burden

The perceived burden of participating in the survey and providing biomeasures is another critical factor. Home visits by nurses or interviewers reduce the effort required from respondents, as they do not need to travel to a clinic, which can be a significant barrier (Clemens et al. 2012). However, interviewers still cannot collect all samples themselves, requiring respondents to take on the burden of the actual biomeasure provision, potentially reducing samples collected. Web surveys can reach a broad audience with minimal intrusion and make it easy for respondents to take part at their convenience, possibly increasing response rates (Bianchi et al. 2017; De Leeuw 2018).

Nonetheless, the absence of human assistance in web surveys may increase the perceived burden when respondents are asked to provide biomeasures, as they might require additional support or reassurance.

Logistical practicalities

The logistical challenges and practical considerations of different survey modes must also be taken into account. Clinic-based or home-visit approaches require substantial resources and can pose accessibility issues for some respondents (Kearney et al. 2011; Clemens et al. 2012). These methods, however, often yield better outcomes for biomeasure data collection (Brown et al. 2019). Mixed-mode designs, such as starting with a web survey followed by an interviewer visit for nonrespondents, aim to combine the strengths of different modes. Web surveys lack the personal touch and real-time reassurance that interviewers can provide, which might be critical when asking for biomeasures. Adding a subsequent face-to-face interaction, if needed, can leverage interviewers' skills in tailoring and persuasion to enhance response rates further and encourage biomeasure provision (Rybak 2023). Mixed-mode surveys offer a cost-effective alternative that can maintain data quality while being adaptable to technological advancements and societal trends. They also provide the flexibility to achieve quicker initial responses, which is a significant operational advantage (De Leeuw 2018).

Data

This paper uses the twelfth wave of the Innovation Panel (IP12), which is part of Understanding Society: The UK Household Longitudinal Study (UKHLS) (University of Essex 2021 University of Essex, Institute for Social and Economic Research 2021). The IP is a vehicle for methodological experimentation in a longitudinal survey design and is conducted annually. The target population is the adult population of Great Britain (England, Scotland, and Wales) 16+ years old. It employs a stratified-clustered design selected through probability proportionate to size methods of persons and households in Great Britain (Lynn 2009). The Innovation Panel includes samples from England, Scotland, and Wales. Postcode sectors from the Postcode Address File were stratified by Government Office Region, the proportion of household heads in National Statistics Socio-Economic Classification (NS-SEC) categories 1 and 2 (nonmanual), and population density. A systematic random sample of 120 postcode sectors primary sampling units (PSU) was taken, with selection probability proportional to population size, measured by number of households. At the fourth (IP4), seventh (IP7), tenth (IP10), and eleventh (IP11) waves, refreshment samples were also drawn using the same 120 PSUs. Waves are conducted annually, and

interviews are attempted with all household members 16 years of age and older. There were 3,693 eligible adults in the selected households comprising the base sample.

We conducted an experiment on the design of a biosocial survey at IP12, where households were randomly and equally allocated to one of three conditions. In one design, a nurse contacted households and attempted to obtain a response. Some respondents in this condition requested to complete online when contacted and would be given log-in information. In the second, the interviewer-first condition, all initial contact was attempted by an interviewer. As with the nurse design, respondents could ask to complete online and were then given log-in details.

In the third condition, the web-first mixed-mode design, households were invited to complete the survey online. After six weeks, web nonrespondents were issued to an interviewer to attempt in-person contact and interviewing. The second and third designs mirror those used in prior waves of Understanding Society. In all conditions, the final three weeks was a "mop-up" stage, where an invitation to take part online was sent to those who had not taken part. Outstanding cases could also be contacted by telephone. Of the 3,693 adults in the base sample, 1,224 were allocated to the web-first design, 1,143 allocated to the interviewer-first design, and 1,328 to the nurse design. All adults in the household were given an unconditional incentive of £10, £20, £30, or £10, with a bonus £20 if everyone in the household completed the web survey within two weeks. Note: this last incentive was not offered to nurse respondents, since they would not be able to complete within the window.

Fieldwork was conducted by Kantar Public and NatCen Social Research and took place July 11 and November 24, 2019. Both fieldwork organizations provided experienced face-to-face interviewers who had worked on Understanding Society in the past. NatCen Social Research provided nurses who were experienced at working on social surveys, such as the Health Survey for England and the English Longitudinal Study of Aging (ELSA). Besides the standard training provided by the fieldwork agencies, there was a study-specific briefing for those working on IP12. Interviewers received a two-day briefing that included reviewing the protocols to collect the health measures. The second day of briefings included accreditations for measuring height, weight, and blood pressure. This involved each interviewer correctly completing the procedures under supervision. Nurses received three days of study-specific training, which included an additional day of briefing on Understanding Society. On the third day, nurses were also accredited to collect the physical measures (Al Baghal et al. 2021).

During the interview, the nurse measured the respondent's height, weight, and blood pressure. Where the respondent was eligible and willing, the nurse also collected a full blood sample. Interviewers collected some biomeasures

in-interview: height, weight, and blood pressure. Those responding to a web interview only provided self-reported biomeasures.

Respondents in all modes were asked to provide dry blood spots (DBS) and a hair sample This can be found in Supplementary Material section A. To test the impact of feedback on outcomes, a random half of the households were offered feedback on the results from the blood sample. Respondents were told they would receive feedback on HDL (high-density lipoprotein), cholesterol, and HbA1c (glycated hemoglobin) levels in their blood samples. Total sample size for each mode design allocation and mode of response by the feedback experiment and incentive structure is presented in Supplementary Material section B.

Nurses took these samples from consenting respondents during the interview. In the other two modes, respondents were asked to complete these independently after the interview. During the interview, the interviewers presented the respondent with a collection kit for DBS and hair, and a leaflet with information. If the respondent was willing to take part in this element of the survey, the interviewer handed over the kit to them. In the web mode, there was a link to an information leaflet online. The respondent was asked whether they would be happy for us to send them the kit, and this kit was then posted to those who consented. For all modes, dried blood spots needed to dry for some time and be packaged and returned by the respondent to the university. In interviewer and web modes, respondents also returned hair samples to the university, while nurses returned hair samples collected. Consent forms were given to all respondents, and signed forms had to be returned to the university with the sample. Respondents returning the kits were sent a £5 thank-you voucher. (For more information about the design and protocol of IP12, see Al Baghal et al. 2021.) The number of respondents providing a fully consented sample (i.e., sample with signed consent form) for each allocated design and mode of response by the feedback experiment and incentive structure is also presented in Supplementary Material section B.

Analysis Methods

Our base sample is the number of adults in issued households, and we use AAPOR RR2 as the outcome (AAPOR 2023), comparing individual respondents completing a full interview to all sampled respondents, including all those in households with unknown eligibility. Those known to be ineligible (e.g., moved out of the country, died) have been removed from the analysis. There were n = 4,251 sampled adults issued at IP12.

The second outcome of interest is a returned biomeasure sample with the consent documentation. Henceforth, we refer to this simply as sample return. The mode of response directly affects sample return, given the importance that channel of communication has on these outcomes (Dykema et al. 2016).

However, unlike the allocated design, the mode of response is not experimental. We therefore analyze this outcome in two ways. We provide an experimental analysis sample return by allocated design, in what is often called an "intention to treat analysis" (ITT). Where compliance with experimental allocation is imperfect, an ITT analysis captures the effect of assignment to a particular condition or treatment, as opposed to the pure effect of experiencing a particular condition. This analysis uses all the issued sample (n = 4,251). Second, we provide an analysis by mode of response, which looks at the effect of the mode in which subjects actually complete the survey. Given that it analyzes the observed mode of response, this analysis is based on the sample of responding adults (n = 2,162).

Dependent variables are all dichotomous: responded to the survey or not, and returned sample or not. Given the necessity of returning the consent form, only samples returned with the form are included in the following analyses; samples without a form included are counted along with other unreturned samples. For all analyses, we analyze the data using PROC SURVEYLOGISTIC in SAS. We account for the stratified-clustered design in all the analyses to estimate correct standard errors; nonresponse weights are not used. There are 60 clusters, and this number is used as the base degrees of freedom for all analyses. We base ITT analyses on the full issued sample, while mode of response is based on respondents. The ability of a respondent to participate in a collection (e.g., hair length) is not measured for interviewer and web modes. The reason a kit was not completed may have been due to inability, but this is not knowable, and hence all noncompletes are treated similarly across all modes.

The number of covariates available for the entire sample is limited, and so only this subset of variables is used in analyses exploring the differences in nonresponse biases across designs. These variables include age and sex of respondent and household location variables. We include whether the respondent lives in an urban or rural area, and what region of Great Britain they live in. There are indicators for North England, England Midlands, Wales, and Scotland, all of which are compared to the south of England (including London and the surrounding areas). There is a limited amount of missing data across these variables, leading to an analytic sample of n = 4,222 for this nonresponse analysis. Where there is missing data, listwise deletion is used in analyses.

For the logistic regression models of sample return, mode of response and the variables in the nonresponse analysis are included, as well as several additional covariates to predict returns of blood and hair samples. These include both survey and respondent characteristics. Time in sample is indicated as whether respondents are in the original sample, samples that are established (IP4 and IP7), or are new samples (IP10 and IP11). Respondents received either incentives of £10, £20, £30, or £10, with a bonus £20 if

everyone in the household completed the web survey within two weeks. Note: this last incentive was not offered to nurse respondents, since they would not be able to complete within the window. A dichotomous indicator is included whether a respondent was offered feedback on their blood sample or not. We include a number of dichotomous indicators for respondent health conditions: cancer, heart disease, stroke, and high blood pressure. Having none of these conditions reported is the baseline. For education of the respondent, those with less than a professional certification are employed as the baseline category compared with those with a professional or university equivalent certification. Supplementary Material section C displays the summary statistics for these variables.

Results

RQ1: Do the Different Approaches Differ in the Response Rate Achieved to the Annual Social Survey?

The top of table 2 shows the result of the experimental design allocation on the survey response rate. The experimental allocation had a noticeable impact on response rates ($F(2, 59) = 3.18$, $p = 0.049$) The post-hoc comparisons of response rates shows that while the web-first design (54.6 percent) outperforms both the interviewer-first (48.2 percent) ($p = 0.019$) and nurse (49.7 percent) ($p = 0.016$)

Table 2. Survey response rate and mode of response by allocated design.

	Allocated design (mode of issue)			
	Nurse	Interviewer first	Web first	Total
Responding	49.7%	48.2%	54.6%	50.6%
Observed N	709	661	792	2,162
$F(2,59) = 3.18$, p $= 0.049$				
Responding mode				
Nurse	95.5%	—-	—-	31.3%
Observed N	677			677
Interviewer	—-	85.0%	26.4%	35.7%
Observed N		562	209	771
Telephone	0.1%	0.6%	0.8%	0.5%
Observed N	1	4	6	11
Web	4.4%	14.4%	72.9%	32.5%
Observed N	31	95	577	703

Note: Response rate is calculated based on mode allocated regardless of mode of response; responding mode shows mode survey completed in within allocations.

in response rates, there is not a detectable difference between the nurse and interviewer-first design ($p = 0.099$).

The lower panel of table 2 shows the mode respondents completed the survey in by allocated design. Most respondents completed the survey in the initially offered mode. The most significant mode switch was in the web-first design, with over a quarter responding via an interviewer. The other two designs both started with an in-person design, and a smaller number of respondents initially approached by an interviewer eventually answered in a different mode than the web-first design. Few of those initially allocated to nurse-led answered in any way other than with a nurse. Only a few cases responded by telephone across designs ($n = 11$), and these are dropped from analyses exploring mode of response.

While the design has some influence on overall response rate, additional analysis can shed light on possible differences in nonresponse bias. As significant research has focused on possible nonresponse bias (e.g., Daikeler et al. 2020), the goal of the following analysis is comparison of possible bias, rather than identifying all the specific bias in the various modes. Table 3 presents logistic regression results predicting survey response in all three designs using available sample characteristics.

The analysis across the nurse-led, interviewer-first, and web-first survey designs suggests a similarity in possible nonresponse bias. None of the location-related variables are significant in any design. In all three modes, age consistently appears as a significant factor, with coefficients indicating an increase in response probability with advancing age. Additionally, new refreshment sample members are significantly less likely to respond in all designs. These similarities suggest that the potential for nonresponse bias is relatively consistent, regardless of the survey design employed.

Sex also shows a similar trend across all three designs. In all designs, females are more likely to respond (all $p < 0.05$), consistent with other findings. This consistent pattern across modes suggests that possible response biases arise in similar ways. The only other difference that arises is that established samples also have a lower response rate than the original sample in the interviewer-first design. Overall, the overarching narrative is that of cross-mode consistency in possible nonresponse bias.

RQ2: How Do the Different Designs Impact the Return of Biological Samples?

Table 4 provides the return rates for the two biomeasures across experimentally allocated designs (accounting for both survey response and sample return) and mode of response (given they have responded to the survey). The ITT analysis of the experiment shows important differences by design in returns of both DBS ($F(2, 59) = 51.83$ $p < 0.0001$) and hair ($F(2,59) = 37.15$ $p < 0.0001$)

Table 3. Logistic regression predicting survey response.

	Allocated design (mode of issue)		
	Nurse Unstandardized coeff. (S.E.; p-value)	Interviewer first Unstandardized coeff. (S.E.; p-value)	Web first Unstandardized coeff. (S.E.; p-value)
Sample (Ref: Original sample)			
Established refreshments	−0.207	−0.360	−0.210
	(0.206; 0.318)	(0.169; 0.037)	(0.191; 0.278)
New refreshments	−0.857	−0.988	−0.923
	(0.196; <0.001)	(0.156; <0.001)	(0.174; <0.001)
Respondent characteristics			
Age	0.022	0.026	0.016
	(0.004; <0.001)	(0.004; <0.001)	(0.004; <0.001)
Female	0.209	0.273	0.307
	(0.094; 0.029)	(0.083; 0.002)	(0.089; 0.009)
Urban	0.069	−0.031	0.077
	(0.190; 0.718)	(0.1864; 0.869)	(0.159; 0.630)
Region (Ref: South England)			
North England	0.040	0.185	−0.093
	(0.195; 0837)	(0.187; 0.325)	(0.225; 0.683)
England Midlands	0.285	0.006	0.023
	(0.263; 0.283)	(0.219; 0.979)	(0.191; 0.906)
Wales	−0.507	0.360	0.420
	(0.0317; 0.115)	(0.590; 0.544)	(0.277; 0.135)
Scotland	−0.303	0.286	−0.311
	(0.219; 0.172)	(0.427; 0.506)	(0.257; 0.231)
Constant	−0.889	−1.044	−0.361
	(0.335; 0.010)	(0.346; 0.004)	(0.309; 0.248)
n	1,413	1,365	1,444
Wald F(9,51)	6.12	9.83	6.88
	(p < 0.001)	(p < 0.001)	(p < 0.001)

Note: Logistic regression predicting response rate (AAPOR RR2) to Innovation Panel 12 of Understanding Society: The UK Household Longitudinal Study by assigned mode.

samples. Post-hoc tests show the difference between the nurse design (35.6 percent, 29.0 percent) and both the web-first (16.3 percent, 13.8 percent) and interviewer-first (20.9 percent, 16.6 percent) designs in terms of sample return for both DBS and hair (interviewer-first DBS $p = 0.041$, hair $p = 0.027$; for web, both $p < 0.0001$) The interviewer-first design appears to lead to better return rates than web first for DBS ($p = 0.006$) and hair ($p = 0.052$). This finding

Table 4. Return rates for the biomeasures by design and mode of response.

Allocated design	Nurse	Interviewer first	Web first	Total
DBS	35.6%	20.9%	16.3%	24.2%
Observed N	508	286	236	1,030
$F(2, 59) = 51.83$ p < 0.0001				
Hair	29.0%	16.6%	13.8%	19.8%
Observed N	414	227	201	842
$F(2, 59) = 37.15$ p < 0.0001				
Mode of response	Nurse	Interviewer	Web	Total
DBS	74.3%	39.8%	31.2%	47.8%
Observed N	503	307	219	1,029
Hair	60.4%	31.8%	26.6%	39.1%
Observed N	409	245	187	841

Note: Return is defined as provision of sample as well as completion of required consent forms.

suggests an important trade-off in designs when collecting survey response and biomeasures. The nurse design obtains the lowest survey response rate but obtains the highest sample return. Conversely, the web-first design performs best for survey response but has the poorest performance for biomeasure sample return.

The bottom panel of table 4 shows sample return by mode of response. In contrast to response rates, nurses (74.3 percent, 60.4 percent) are much better than other modes in obtaining samples from respondents. Interviewers (39.8 percent, 31.8 percent), while falling short of what nurses achieved, did outperform the web (31.2 percent, 26.6 percent) for both DBS and hair. For all the allocated designs and modes of response, sample return was also higher for DBS than for hair.

Logistic regression models are used to explore the impact of mode of response on DBS and hair sample return. The models control for survey factors and respondent characteristics using the covariates described above. Table 5 presents the results in terms of odds ratios. Even after controlling for survey factors and respondent characteristics, the impact of mode of response remains. Compared to nurses, interviewer and web respondents are less likely to provide either DBS or hair samples. The results imply that nurses are 4.59 times more likely to obtain a DBS sample than an interviewer and 6.71 times more likely than from a web respondent. Similarly, nurses are 3.72 more likely than interviewers and 4.46 times more likely than from web respondents to obtain a hair sample. Coefficient contrasts show that interviewers' estimated likelihood of

Table 5. Logistic regression predicting biological sample return.

	DBS return Unstandardized coeff. (S.E.; *p*-value)	Hair return Unstandardized coeff. (S.E.; *p*-value)
Mode (Ref: Nurse)		
Interviewer	−1.525	−1.341
	(0.152; <0.001)	(0.145; <0.001)
Web	−1.901	−1.500
	(0.133; <0.001)	(0.145; <0.001)
Sample (Ref: Original sample)		
Established refreshments	−0.0002	−0.145
	(0.127; 0.999)	(0.141; 0.308)
New refreshments	0.012	−0.083
	(0.158; 0.943)	(0.147; 0.578)
Incentive (Ref: £10)		
£20	0.182	0.270
	(0.256; 0.481)	(0.271; 0.324)
£30	0.125	0.263
	(0.128; 0.334)	(0.147; 0.079)
£10+£20	−0.149	−0.102
	(0.261; 0.57)	(0.281; 0.718)
Provide feedback	0.280	0.416
	(0.115; 0.018)	(0.106; <0.001)
Health conditions		
Cancer	0.066	0.113
	(0.239; 0.785)	(0.222; 0.611)
Heart disease	−0.032	0.420
	(0.236; 0.892)	(0.234; 0.078)
Stroke	−1.425	−0.920
	(0.453; 0.003)	(0.354; 0.012)
High BP	0.332	0.094
	(0.136; 0.018)	(0.134; 0.486)
Diabetes	0.243	0.182
	(0.174; 0.168)	(0.191; 0.344)
Respondent characteristics		
Age	0.012	0.018
	(0.003; 0.001)	(0.004; <0.001)
Female	0.145	0.875
	(0.096; 0.136)	(0.101; <0.001)
Urban	0.031	−0.062
	(0.077; 0.69)	(0.079; 0.431)

Table 5. Continued.

	DBS return Unstandardized coeff. (S.E.; *p*-value)	Hair return Unstandardized coeff. (S.E.; *p*-value)
Region (Ref: South England)		
North England	0.055	−0.286
	(0.159; 0.731)	(0.167; 0.091)
England Midlands	0.054	−0.200
	(0.203; 0.792)	(0.143; 0.169)
Wales	−0.074	0.190
	(0.337; 0.826)	(0.266; 0.476)
Scotland	0.152	−0.231
	(0.236; 0.522)	(0.269; 0.395)
Education (Ref: Less than A-level)		
University degree	0.477	0.290
	(0.125; 0.0)	(0.115; 0.014)
Professional/A-level	0.187	0.101
	(0.133; 0.162)	(0.139; 0.473)
Constant	−0.073	−1.200
	(0.246; 0.769)	(0.295; <0.001)
Wald F(22,39)	10.34	9.80
	(*p* < 0.001)	(*p* < 0.001)

Note: Logistic regression predicting return rates of biological samples in IP12 among survey respondents. Mode estimates effect of mode on response.
n = 2,105.

obtaining a DBS return is larger than from web respondents ($p = 0.009$). However, the estimated effects for interviewer and web are not clearly different for the return of a hair sample ($p = 0.294$).

Beyond mode of response, the only other survey factor that has a significant impact is the provision of feedback. Respondents are significantly more likely to provide both biological samples if offered feedback. These results are consistent with other analyses of this data (Benzeval et al. 2023). Even though feedback was specifically related to the blood sample, the offer also led to a significant increase in the provision of hair samples. Other survey factors, such as time in sample and incentive amount, are not significantly related to return of either sample.

As in the nonresponse analysis, regional variables have no detectable impact on sample return. There is some impact of health conditions on sample return. Respondents that have had a stroke are less likely to provide both samples; whether this is related to any motor function limitations is unclear.

Those with high blood pressure are more likely to return a DBS sample, but there is no identified effect for the return of a hair sample. Unlike Sakshaug et al. (2010), those diagnosed with diabetes were no more likely to provide measures than undiagnosed respondents. That there are few identified conditions with somewhat inconsistent effects is consistent with other studies (e.g., Sakshaug et al. 2010), where respondents with underlying health conditions are generally not any more likely to provide a sample than those without, with a few exceptions.

Results show different respondent effects for provision of DBS and hair samples. Older respondents and those with more education are more likely to provide both biological samples. Several studies have found a similar impact of education (Jaszczak et al. 2009; Sakshaug et al. 2010; Dykema et al. 2016; Cernat et al. 2021); these studies have mostly used studies of older adults, and fewer have found a similar impact of age (Cernat et al. 2021). Females are more likely to provide a hair sample than males, with no little detectable effect of sex on DBS provision. The difference in outcomes is likely in part due to differences in the amount of hair available. While measures are not available for interviewer and web respondents, nurses observed and indicated if a respondent had three centimeters or more of hair. More than double the percentage of females (78.4 percent) than males (38.6 percent) had this length of hair.

Given the impacts of design and mode on both response and sample return, costs should be explored to identify the costs and benefits. The costs provided by the fieldwork agency are considered proprietary, and to study costs, we explore relative costs between the three designs. We use the nurse design as the baseline for cost per achieved interview and DBS sample (hair sample is the same pattern), and provide the percentage cost for the other two designs in regard to these two outcomes. Table 6 presents these relative costs.

Overall, nurses represent the highest expense. This mode, however, offers qualitative benefits due to professional expertise, which can be vital when particular biological samples are needed and/or the integrity of samples is critical. The web-first mode, while being the most cost-effective for survey

Table 6. Relative costs per interview and sample returned by design.

	Nurse	Interviewer first	Web first
Cost per interview	100%	71.3%	44.9%
Cost per sample	100%	109.5%	92.7%

Note: Nurse design is the baseline cost, and presented at 100 percent. Other costs relative to nurse design.

responses at 44.9 percent of the nurse design's cost, sees a substantial increase in relative cost for DBS sample collection at 92.7 percent. This suggests a lower efficiency in translating survey responses to DBS samples, highlighting the trade-offs between initial survey costs and the effectiveness of biological data collection.

The interviewer-first mode, which costs 71.3 percent of the nurse mode for survey responses and 109.5 percent for DBS sample returns, offers a compelling balance. It is more economical than the nurse mode for survey collection, but slightly more expensive for DBS sample returns, implying a higher cost efficiency in converting survey responses to DBS samples. This mode is particularly advantageous for studies that prioritize the collection of biological data without heavily compromising on cost. The strategic selection of a survey mode thus hinges on a delicate balance between financial constraints and the specific goals of the survey. For maximizing efficiency in DBS sample collection, the nurse mode is best, despite its higher cost. Conversely, for surveys where cost is a critical factor, the web-first mode is the most cost-effective for gathering responses, though with a trade-off in DBS sample collection efficiency. The interviewer-first design offers a mix of cost efficiency and a high conversion rate of survey responses to DBS samples.

Discussion

Collecting biomeasures as part of a survey design can add new understanding to both sources of data. This study is the first to compare a nurse led design to interviewer-first and web-first mixed-mode designs. In the nurse led, nurses obtain response and collect biomeasures, while in interviewer first and web first, response is obtained as in a standard survey, but biomeasures are collected through provision of a kit. Recent findings suggest little measurement differences in biomeasures when collected by nurses or self-collected by respondents (Kumari et al. 2023), suggesting the importance of design is mainly related to obtaining response and sample return.

We find that the nurse design is not significantly worse in obtaining survey response than the interviewer-first design, contrary to previous findings (Brown et al. 2019). The web-first design performed significantly better than other designs in obtaining survey response, which has been found in some other studies (De Leeuw 2018). Provision of the web option appears to be key, given the high take-up of this option. However, just over a quarter in the web-first design did not respond until an interviewer visited the house, indicating the key role interviewers play in increasing response rates.

For the return of biomeasures, the nurse design is clearly the most effective. Nurses may be more trusted by respondents to collect biomeasures, as this is seen to be more closely related to their job. Also, providing DBS and

hair samples was done at the point of interview by the nurse, although return of the DBS in the mail relied on the respondent. Conversely, for interviewer and web respondents, the kit was handed over to the respondent or sent to them through the post. The respondent was therefore responsible for both taking the sample and then posting it back afterward.

The survey response and sample return rates show a design trade-off in design between obtaining interviews and returned samples. Those who are invited to complete their interview online first are more likely to respond, but web respondents are less likely to return a biomeasure sample. Those issued to nurses are less likely to respond to the survey, but most likely to return a biomeasure sample. If biomeasures are the key outcome, and if it is financially and logistically viable, a nurse design may be best. However, in a context where biomeasures are of equal importance to the social survey, or it is not feasible or financially viable to issue a large sample to nurses, a judicious balance between interviewers and web self-completion is required to maximize both the response to the interview and the return of biomeasures. Providing feedback also appears key in the provision of samples, with other analyses of these data showing this especially increases return in the interviewer and web modes (Benzeval et al. 2023).

While the finding identifies useful methodological considerations for collecting survey and biomeasure data together, the generalizability of these findings to cross-sectional surveys warrants consideration. While longitudinal surveys benefit from established respondent familiarity and engagement, cross-sectional surveys lack this continuity, potentially impacting response rates and sample returns differently. This is indicated in the lower survey response rate for newer samples relative to more established samples across all the designs. In the context of cross-sectional surveys, the effectiveness of the mixed-mode designs might vary due to the absence of longitudinal rapport, and the initial engagement strategies might need adjustment.

Other models of biomeasure collection could be employed that are not tested here. We did not explore nurses with an interviewer or web follow-up mixed-mode design, for example, nor a web-first design for the survey completion with a nurse or interviewer follow-up to collect biomeasures. The distinct nature of biomeasure sample collection in a survey environment poses unique challenges that might not be entirely captured in this study, such as respondent privacy concerns and logistical complexities. Future studies may also need to consider whether the expanded use of incentives, for example by offering a larger incentive on return of a sample, or a more thorough use of reminders for those who have received a kit through the post, may increase the sample return rate. Qualitative research to understand the dynamics behind why people may or may not want to provide samples would also be informative, as has been done for other biomeasure sample types (Attah et al. 2023).

Supplementary Material

Funding

Data Availability

Replication data and documentation are available at https://osf.io/gdj2m/ and http://doi.org/10.5255/UKDA-SN-6849-15.

References

AAPOR (American Association for Public Opinion Research). (2023). Standard definitions: Final dispositions of case codes and outcome rates for surveys. 10th ed. https://aapor.org/wp-content/uploads/2023/05/Standards-Definitions-10th-edition.pdf.

Al Baghal, Tarek, Michaela Benzeval, Jonathan Burton, Thomas F. Crossley, Meena Kumari, and Shavanthi Rajatileka. 2021. "Collection of biomarkers using nurses, interviewers, and participants: the design of IP12." Understanding Society Working Paper 2021-06. University of Essex, Colchester. https://www.iser.essex.ac.uk/research/publications/working-papers/understanding-society/2021-06.

Allen, Caitlin G., Kelly J. Hunt, Lori L. McMahon, Clay Thornhill, Amy Jackson, John T. Clark, Katie Kirchoff, Kelli L. Garrison, Kimberly Foil, Libby Malphrus, Samantha Norman, Paula S. Ramos, Kelly Perritt, Caroline Brown, Leslie Lenert, and Daniel P. Judge. 2024. "Using Implementation Science to Evaluate a Population-Wide Genomic Screening Program: Findings from the First 20,000 In Our DNA SC Participants." *American Journal of Human Genetics* 111:433–44. https://doi.org/10.1016/j.ajhg.2024.01.004.

Attah, Ramlatu, Katy Lawn, Avindri Chandraharan, Violetta Parutis, and Jonathan Burton. 2023. "Understanding Society Microbiome Study: findings from qualitative research into the clarity, completeness and accessibility of prospective participant materials for the Wave 16 microbiome sample collection." Understanding Society Working Paper 2023-14, University of Essex, Colchester. https://www.understandingsociety.ac.uk/research/publications/568027.

Avendano, Mauricio, Annette Scherpenzeel, and Johan P. Mackenbach. 2010. "Can biomarkers be collected in an internet survey? A pilot study in the LISS panel." In: *Social and Behavioral Research and the Internet: Advances in Applied Methods and Research Strategies. European association of methodology series*, edited by Marcel Das, Peter Ester, and Lars Kaczmirek, 371–409. London, UK: Routledge.

Benzeval, Michaela, Alexandria Andrayas, Jan Mazza, Tarek Al Baghal, Jonathon Burton, Thomas F. Crossley, and Meena Kumari. 2023. "Does the Feedback of Blood Results in

Observational Studies Influence Response and Consent? A Randomised Study of the Understanding Society Innovation Panel." *BMC Medical Research Methodology* 23:134. https://doi.org/10.1186/s12874-023-01948-y.

Bianchi, Annamaria, Meena Kumari, and Andrew M. Jones. 2016. "How Do Biomarkers and Genetics Contribute to Understanding Society?" *Health Economics* 25:1219–22. https://doi.org/10.1002/hec.3400.

Bianchi, Annamaria, Silvia Biffignandi, and Peter Lynn. 2017. "Web-Face-to-Face Mixed-Mode Design in a Longitudinal Survey: Effects on Participation Rates, Sample Composition, and Costs." *Journal of Official Statistics* 33:385–408. https://doi.org/10.1515/jos-2017-0019.

Bowling, Ann. 2005. "Mode of Questionnaire Administration Can Have Serious Effects on Data Quality." *Journal of Public Health (Oxford, England)* 27:281–91. https://doi.org/10.1093/pubmed/fdi031.

Brown, Matt, Emily Gilbert, Lisa Calderwood, Kate Taylor, and Hannah Morgan. 2019. "Collecting Biomedical and Social Data in a Longitudinal Survey: A Comparison of Two Approaches." *Longitudinal and Life Course Studies* 10:453.

Burton, Jonathan, Mick P. Couper, and Annette Jäckle. 2024. "The Effects of Placement and Order on Consent to Data Linkage in a Web Survey." *Journal of Survey Statistics and Methodology* 12:1212–23. https://doi.org/10.1093/jssam/smae004.

Clemens, Samantha, Lisa Given, and Susan Purdon. 2012. "Methods of collecting biological data: considerations, challenges and implications." Presentation to the 67th annual American Association for Public Opinion Research Conference. Orlando, FL.

Cernat, Alexandru, Joseph W. Sakshaug, Tarani Chandola, James Nazroo, and Natalie Shlomo. 2021. "Nurse Effects on Non-Response in Survey-Based Biomeasures." *International Journal of Social Research Methodology* 24:487–99. https://doi.org/10.1080/13645579.2020.1832737.

Chartier, Karen G., Priscilla Martinez, Cory Cummings, Brien P. Riley, and Katherine J. Karriker-Jaffe. 2021. "Recruiting for Diversity: A Pilot Test of Recruitment Strategies for a National Alcohol Survey with Mail-in Genetic Data Collection." *Journal of Community Genetics* 12:459–68. https://doi.org/10.1007/s12687-020-00502-3.

Daikeler Jessica, Michael Bošnjak, and Katja Lozar Manfreda. 2020. "Web Versus Other Survey Modes: An Updated and Extended Meta-Analysis Comparing Response Rates." *Journal of Survey Statistics and Methodology* 8:513–39. https://doi.org/10.1093/jssam/smz008.

De Leeuw, Edith D. 2005. "To Mix or Not to Mix Data Collection Modes in Surveys." *Journal of Official Statistics* 21:233–55.

——— 2018. "Mixed-Mode: Past, Present, and Future." *Survey Research Methods* 12:75–89. https://doi.org/10.18148/SRM/2018.V12I2.7402.

Dykema, Jennifer, Kerryann DiLoreto, Kenneth D. Croes, Dana Garbarski, and Jeremy Beach. 2016. "Factors Associated with Participation in the Collection of Saliva Samples by Mail in a Survey of Older Adults." *Public Opinion Quarterly* nfw045. https://doi.org/10.1093/poq/nfw045.

Fuller, Garth, Kelly Njune Mouapi, Sandy Joung, Chrisandra Shufelt, Irene van den Broek, Mayra Lopez, Aaron Robinson, Shivani Dhawan, Mitra Mastali, Brennan Spiegel, Noel Bairey Merz, and Jennifer E. Van Eyk. 2019. "Feasibility of Patient-Centric Remote Dried Blood Sampling: The Prediction, Risk, and Evaluation of Major Adverse Cardiac Events (PRE-MACE) Study." *Biodemography and Social Biology* 65:313–22. https://doi.org/10.1080/19485565.2020.1765735.

Gatny, Heather H., Mick P. Couper, and William G. Axinn. 2013. "New strategies for Biosample Collection in Population-based Social Research" *Social Science Research* 42:1402–9. https://doi.org/10.1016/j.ssresearch.2013.03.004.

Ghafari, Mahan, Matthew Hall, Tanya Golubchik, Daniel Ayoubkhani, Thomas House, George MacIntyre-Cockett, Helen R. Fryer, Laura Thomson, Anel Nurtay, Steven A. Kemp, Luca

Ferretti, David Buck, Angie Green, Amy Trebes, Paolo Piazza, Lorne J. Lonie, Ruth Studley, Emma Rourke, Darren L. Smith, Matthew Bashton, Andrew Nelson, Matthew Crown, Clare McCann, Gregory R. Young, Rui Andre Nunes Dos Santos, Zack Richards, Mohammad Adnan Tariq, Roberto Cahuantzi, Jeff Barrett, Christophe Fraser, David Bonsall, Ann Sarah Walker, and Katrina Lythgoe; COVID-19 Genomics UK (COG-UK) Consortium. 2024. "Prevalence of Persistent SARS-CoV-2 in a Large Community Surveillance Study." *Nature* 626:1094–101. https://doi.org/10.1038/s41586-024-07029-4.

Golding, Jean. and ALSPAC Study Team. 2004. "The Avon Longitudinal Study of Parents and Children (ALSPAC)—Study Design and Collaborative Opportunities." *European Journal of Endocrinology* 151 Suppl 3: U119–23. https://doi.org/10.1530/eje.0.151u119.

Guyer, Heidi, Mary Beth Ofstedal, Carli Lessof, and Kate Cox. 2017. *The benefits and challenges of collecting physical measures and biomarkers in cross-national studies*. Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan, https://hrs.isr.umich.edu/sites/default/files/biblio/Collecting%20PM-Bio%20Data_DocumentationReport.pdf.

Jäckle, Annette, Kelsey Beninger, Jonathanm Burton, and Mick P. Couper. 2021. "Understanding Data Linkage Consent in Longitudinal Surveys." In *Advances in Longitudinal Survey Methodology*, edited by Peter Lynn, 122–50. Wiley. https://doi.org/10.1002/9781119376965.

Jaszczak, Angela, Katie Lundeen, and Stephen Smith. 2009. "Using Nonmedically Trained Interviewers to Collect Biomeasuresin a National In-Home Survey." *Field Methods* 21:26–48. https://doi.org/10.1177/1525822X08323988.

Kearney, Patricia M., Hilary Cronin, Claire O'Regan, Yumiko Kamiya, Brendan J. Whelan, and Rose Anne Kenny. 2011. "Comparison of Centre and Home-Based Health Assessments: Early Experience from the Irish Longitudinal Study on Ageing (TILDA)." *Age and Ageing* 40:85–90. https://doi.org/10.1093/ageing/afq124.

Kenny, Rose Anne, Brendan J. Whelan, Hilary Cronin, Yumiko Kamiya, Patricia Kearney, Claire O'Regan, Mélina Ziegel. 2010. *The Design of the Irish Longitudinal Study on Ageing*. Dublin: Trinity College. https://tilda.tcd.ie/publications/reports/pdf/Report_DesignReport.pdf.

Korbmacher, Julie M., and Mathis Schroeder. 2013. "Consent When Linking Survey Data with Administrative Records: The Role of the Interviewer." *Survey Research Methods* 7:115–31. https://doi.org/10.18148/SRM/2013.V7I2.5067.

Kuh, Diana, Mary Pierce, Judith Adams, John Deanfield, Ulf Ekelund, Peter Friberg, Arjun K Ghosh, Nikki Harwood, Alun Hughes, Peter W Macfarlane, Gita Mishra, Denis Pellerin, Andrew Wong, Alison M Stephen, Marcus Richards, and Rebecca Hardy; NSHD Scientific and Data Collection Team. 2011. "Cohort Profile: Updating the Cohort Profile for the MRC National Survey of Health and Development: A New Clinic-Based Data Collection for Ageing Research." *International Journal of Epidemiology* 40:e1–e9. https://doi.org/10.1093/ije/dyq231.

Kumari, Meena, Alexandria Andrayas, Tarek Al Baghal, Jonathan Burton, Thomas F. Crossley, Kerry S. Jones, Damon A. Parkington, Albert Koulman, and Michaela Benzeval. 2023. "A Randomised Study of Nurse Collected Venous Blood and Self-Collected Dried Blood Spots for the Assessment of Cardiovascular Risk Factors in the Understanding Society Innovation Panel." *Scientific Reports* 13:13008. https://doi.org/10.1038/s41598-023-39674-6.

Lynn, P. 2009. "Sample design for Understanding Society", Understanding Society Working Paper, 2009-01. https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2009-01.pdf.

Marmot, Michael, and Eric Brunner. 2005. "Cohort Profile: The Whitehall II Study." *International Journal of Epidemiology* 34:251–56. https://doi.org/10.1093/ije/dyh372.

Marmot, Michael and Andrew Steptoe. 2008. "Whitehall II and ELSA: Integrating Epidemiological and Psychobiological Approaches to the Assessment of Biological Indicators." In *Biosocial Surveys*, edited by Weinstein Maxine, James W. Vaupel and

Kenneth W. Wachter, 42–59. Washington, DC: National Academy Press. https://www.ncbi.nlm.nih.gov/books/NBK62436/.

McFall, Stephanie L., Cara Booker, Jonathan Burton, and Anne Conolly. 2012. Implementing the biosocial component of Understanding Society—nurse collection of biomeasures. Understanding Society Working Paper, 2012-04. https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2012-04.pdf.

Mindell, Jennifer, Jane P. Biddulph, Vasant Hirani, Emanuel Stamatakis, Rachel Craig, Susan Nunn, and Nicola Shelton. 2012. "Cohort Profile: The Health Survey for England." *International Journal of Epidemiology* 41:1585–93. https://doi.org/10.1093/ije/dyr199.

Olson, Kristen, Jolene D. Smyth, Rachel Horwitz, Scott Keeter, Virginia Lesser, Stephanie Marken, Nancy A. Mathiowetz, Jaki S. McCarthy, Eileen O'Brien, Jean D. Opsomer, Darby Steiger, David Sterrett, Jennifer Su, Z. Tuba Suzer-Gurtekin, Chintan Turakhia, and James Wagner. 2021. "Transitions from Telephone Surveys to Self-Administered and Mixed-Mode Surveys: AAPOR Task Force Report." *Journal of Survey Statistics and Methodology* 9:381–411. https://doi.org/10.1093/jssam/smz062.

Peakman, Tim C., and Paul Elliott. 2008. "The UK Biobank Sample Handling and Storage Validation Studies." *International Journal of Epidemiology* 37 Suppl 1:i2–i6. https://doi.org/10.1093/ije/dyn019.

Riley, Steven, Christina Atchison, Deborah Ashby, Christl A. Donnelly, Wendy Barclay, Graham S. Cooke, Helen Ward, Ara Darzi, and Paul Elliott; REACT Study Group. 2021. "REal-Time Assessment of Community Transmission (REACT) of SARS-CoV-2 Virus: Study Protocol." *Wellcome Open Research* 5:200. https://doi.org/10.12688/wellcomeopenres.16228.2.

Rybak, Adam. 2023. "Survey Mode and Nonresponse Bias: A Meta-Analysis Based on the Data from the International Social Survey Programme Waves 1996–2018 and the European Social Survey Rounds 1 to 9." *PLoS ONE* 18:e0283092. https://doi.org/10.1371/journal.pone.0283092.

Sakshaug, Joseph W, Alexandru Cernat, and Trivellore E. Raghunathan 2019. "Do Sequential Mixed-Mode Surveys Decrease Nonresponse Bias, Measurement Error Bias, and Total Bias? An Experimental Study." *Journal of Survey Statistics and Methodology* 7:545–71. https://doi.org/10.1093/jssam/smy024.

Sakshaug, Joseph W., Mick P. Couper, and Mary Beth Ofstedal. 2010. "Characteristics of Physical Measurement Consent in a Population-Based Survey of Older Adults." *Medical Care* 48:64–71. https://doi.org/10.1097/MLR.0b013e3181adcbd3.

Sakshaug, Joseph W., Sebastian Hülle, Alexandra Schmucker, and Stefan Liebig. 2017. "Exploring the Effects of Interviewer- and Self-Administered Survey Modes on Record Linkage Consent Rates and Bias." *Survey Research Methods* 11:171–88. https://doi.org/10.18148/SRM/2017.V11I2.7158.

Sonnega, Amanda, Jessica D. Faul, Mary Beth Ofstedal, Kenneth M. Langa, John. W.R. Phillips, and David R. Weir. 2014. "Cohort Profile: The Health and Retirement Study (HRS)." *International Journal of Epidemiology* 43:576–85. https://doi.org/10.1093/ije/dyu067.

Steptoe, Andrew, Elizabeth Breeze, James Banks, and James Nazroo. 2013. "Cohort Profile: The English Longitudinal Study of Ageing." *International Journal of Epidemiology* 42:1640–48. https://doi.org/10.1093/ije/dys168.

UK Biobank. 2007. "Protocol for a large-scale prospective epidemiological resource, Protocol for a large-scale prospective epidemiological resource." https://www.ukbiobank.ac.uk/media/gnkeyh2q/study-rationale.pdf.

University of Essex, Institute for Social and Economic Research. 2021. *Understanding Society: Innovation Panel, Waves 1-13, 2008-2020*. [Data collection]. 11th ed. UK Data Service. SN: 6849. http://doi.org/10.5255/UKDA-SN-6849-14.

Weir, David. 2008. Elastic powers: The integration of biomarkers into the Health and Retirement Study. In *Biosocial Surveys*, edited by Maxine Weinstein, James W. Vaupel, and Kenneth W. Wachter, 78–95. Washington, DC: National Academy Press. https://www.ncbi.nlm.nih.gov/books/NBK62436/.

Weiss, Luzia M., Joseph W. Sakshaug, and Axel Borsch-Supan. 2019. "Collection of biomeasur-esin a cross-national setting: Experiences in SHARE". In *Advances in comparative survey methods: Multinational, Multiregional and Multicultural Contexts (3MC)*, edited by Timothy Johnson, Pennell Beth-Ellen, Stoop Ineke, & Dorer Brita, 623–41. John Wiley and Sons.

Zipf, George, Michele Chiappa, Kathryn S. Porter, Yechiam Ostchega, Brenda G. Lewis, and Jennifer Dostal. 2013. "National Health and Nutrition Examination Survey: Plan and Operations, 1999-2010." *Vital and Health Statistics. Ser. 1, Programs and Collection Procedures* 56:1–37. no(August).