Collecting Hair Samples in Online Panel Surveys: Participation Rates, Selective Participation, and Effects on Attrition

Mario Lawes[1] · Clemens Hetschko[1,2,3] · Joseph W. Sakshaug[4,5,6] · Michael Eid[1]

[1]Freie Universität Berlin
[2]University of Leeds
[3]CESifo
[4]Institute for Employment Research
[5]LMU Munich
[6]University of Mannheim

Combining survey data with biological information allows examining complex interrelationships between a person's physiological status and behavioral or health-related outcomes. Given the increasing importance of online surveys and smartphone-based research, a crucial question is whether biomarker collection can be embedded in online surveys without any face-to-face interaction. The present study addresses this question and investigated participation rates and selective participation in a longitudinal hair collection study that was embedded within an app-based smartphone panel survey on the well-being of German jobseekers. The study further examined the association between participating in the first hair collection wave and panel attrition. The results indicate that the vast majority (81%) of individuals was willing to participate in the first hair collection wave with only a few selection effects. Only older age and higher levels of perceived stress were modestly associated with the stated willingness to participate in the first hair collection wave. The strongest selectivity was induced by the inevitable exclusion of individuals with short hair styles, which led to an underrepresentation of men. Furthermore, respondents' purported willingness to participate in the first hair collection wave and their actual participation was largely disconnected. This lack of compliance decreased in subsequent collection waves. Notably, participating in the first hair collection wave was positively related to long-term panel participation. Overall, the study underlines the general feasibility of integrating biomarker collections into online surveys.

Keywords: biomarker collection; hair cortisol; online surveys; selective participation; panel attrition

1   Introduction

Across the social and behavioral sciences, there is widespread interest in the measurement of physical and mental health. Often, questionnaire studies are the method of choice, as they allow measuring health in an economical manner for both respondents and researchers. While the subjective nature of answers to questionnaire items is vital for assessing a person's mental well-being or experience of pain, self-reports are limited when it comes to the measurement of physical health. For this purpose, the collection of bio-

logical measurements (hereafter referred to as biomarkers) alongside questionnaire data has increased immensely in recent times (Boyle et al., 2021; Sakshaug et al., 2015). Compared to questionnaire data, biomarkers are less prone to self-report distortions (e.g., social desirability bias), and can point to medical conditions respondents are unaware of. Combining questionnaire and biomarker data allows for a better understanding of the complex interrelationships between a person's physiological status, physical and social environment, behaviors, and health outcomes (Weinstein et al., 2008).

While biomarkers have long been collected in clinical studies, new assessment tools have made it possible to also collect them in population-based field studies. For instance, biomarker data were collected within add-on studies of

Corresponding author: Mario Lawes, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany (Email: mario.lawes@fu-berlin.de)

large-scale panel surveys such as the US Health and Retirement Study (HRS; Juster & Suzman, 1995), the UK Household Longitudinal Study (UKHLS; University of Essex, Institute for Social and Economic Research, 2022), the Survey of Health and Retirement in Europe (SHARE; Börsch-Supan et al., 2005), and the Dutch Longitudinal Internet studies for the Social Sciences (LISS; Avendano et al., 2011). Moreover, many biomarkers can be self-collected by study participants without the assistance of interviewers or nurses, which decreases collection costs and time. For instance, hair samples can be easily self-collected in the homes of the study participants. Such hair samples may, for example, be used to determine hair cortisol concentration (HCC), which is a reliable biomarker of chronic stress that is strongly linked to ill-health (Chrousos, 2009; Heuser et al., 2003; Tafet & Nemeroff, 2016). The analysis of HCC has gained increasing interest in the social and behavioral sciences (Greff et al., 2019; Kirschbaum et al., 2009; Lawes, Hetschko, Schöb, et al., 2022; Penz et al., 2018; Schaafsma et al., 2021).

When biomarker data can be obtained without any face-to-face contact with respondents, it might also be possible to integrate biomarker collection into online surveys. Here, an important distinction needs to be made between smartphone-integrated biosensors that are automatically collected (Perry et al., 2018) and smartphone-based surveys that ask respondents to manually collect the biomarker measurements. In the present article, we focus on the latter. Online surveys play an increasingly important role in the social sciences as they can reduce response burden for participants, allow convenient high-frequency measurements at low cost, and can feature interventions, push-notifications, and incentives (Miller, 2012). Integrating biomarker collection into online studies also makes it possible to link biomarker data with other innovative data sources. For example, in smartphone-based studies biomarker data can be linked to geolocation, movement patterns, browsing history, communication behavior (Keusch et al., 2019), or momentary mood assessments via the experience sampling method (Hektner et al., 2007; OECD, 2013).

While the self-collection of biomarker material in online surveys likely (i) lowers the cost of data collection, (ii) offers the participants more flexibility and (iii) promises novel data linkages, it also raises several concerns. First, the self-collection of biomarker material in online surveys might be too complicated for participants, resulting in low biomarker participation rates (see Avendano et al., 2011). Second, participation in the biomarker assessment could be highly selective, which has substantial implications for the analysis and the validity of the study conclusions. Third, concerns arise in online panel studies when biomarker data are collected over multiple time points. For instance, the effects of participation burden can accumulate, which might po-

tentially increase dropout rates (Pashazadeh et al., 2021). The present study examines the severity of these issues in an add-on HCC study that was embedded in an app-based smartphone panel survey without any face-to-face interaction with its respondents. We begin by documenting past research that has examined the described concerns with integrating biomarker collection within online surveys.

2   Background

2.1   Participation Rates in (Self-Collection) Biomarker Studies

The first concern of integrating biomarker collection within online surveys is that individuals might perceive the self-collection of biomarker material as burdensome, unpleasant, and intrusive, which could lead to low participation rates. This effect might even be amplified when the biomarker collection is integrated in app-based surveys as these are also known to yield low response rates (for examples, see Hetschko et al., 2022; Jäckle et al., 2019; McGeeney & Weisel, 2015). However, there is a lack of empirical studies examining participation rates of self-collected biomarkers that are integrated within app-based surveys without face-to-face interaction. Initial evidence comes from studies that have compared biomarker participation rates with and without self-collection. These studies indicate that surveys that involve self-collection of biomarkers generally have lower biomarker participation rates than those in which the biomarker is collected with the assistance of a nurse or interviewer (Gatny et al., 2013; Sakshaug et al., 2015). Avendano et al. (2011) reported that only about 19% of invited individuals of a small pilot study based on the LISS panel sent in self-collected blood samples while 15% sent in self-collected saliva samples for cortisol measurements. Similarly low participation rates were found in a study by Etter and Bullen (2011), who reported that only 16% of invited study participants sent in a self-collected saliva sample. Somewhat higher participation rates were reported in a study by Gatny et al. (2013), in which 65% of the invited individuals sent in a self-collected saliva sample. Further evidence comes from studies that investigated the extent to which panelists would be willing to perform additional tasks in panel surveys (Boyle et al., 2021; Revilla et al., 2019). These studies underline that only a small fraction of panelists reported that they would be willing to provide biomarker data (e.g., saliva sample, blood pressure measurements) in addition to their survey responses. Crucially, as these studies are based on hypothetical willingness to perform additional tasks, they likely still overestimate *actual* participation rates in biomarker

collections as some individuals will generally not provide their biomarker data even though they initially reported that they would be willing to do so. Overall, the existing research has therefore indicated that biomarker participation rates in population-based studies are low, particularly, when the biomarker material has to be self-collected by the respondents. To the best of our knowledge there are no studies that have examined biomarker participation rates in self-collection studies that are embedded in online surveys.

2.2 Selective Participation in (Self-Collection) Biomarker Studies

The second concern is that biomarker participation in self-collection biomarker studies could be highly selective putting the validity of the analysis at risk. However, research directly examining this issue is currently lacking. Related insights come from several studies that have investigated patterns of selective participation in biomarker studies in general. For example, Sakshaug et al. (2010) found that Hispanics, people with low confidentiality concerns and, notably, those with at least one doctor visit over the last year were more likely to consent to the collection of biomarkers in the HRS. Schonlau et al. (2010) further reported that risk aversion and being male was associated with a reduced willingness to participate in the collection of buccal cells via cheek swabs and mouth wash from pre-test participants of the German Socio-Economic Panel Study (SOEP). Dykema et al. (2017) found that over 50% of respondents of a mail-based survey as part of a separate data collection in the Wisconsin Longitudinal Study (WLS) provided a saliva DNA sample. In their multivariate analyses, compliance was lower for females, those with lower cognitive ability, and those whose past behavior indicated resistance to survey participation. Gatny et al. (2013) also found that below-average prior cooperation was associated with a lower likelihood of complying with the request to self-collect and mail-back a saliva sample. Moreover, Ford et al. (2016) reported that Whites were more likely than Blacks to participate in a combined ecological momentary assessment and hair cortisol study conducted via face-to-face visits. What is more, men are generally underrepresented in hair collection studies as individuals with short scalp hair cannot be considered (for reviews, see Russell et al., 2012; Stalder & Kirschbaum, 2012). This underlines the importance of differentiating between selective participation due to individuals' preferences to take part in the biomarker collection and eligibility criteria that must be met in order to participate (e.g., availability of at least 2 cm strands of scalp hair for the analysis of hair cortisol).

The respondent's personality traits might also influence participation in biomarker studies as they have been linked with study commitment (see e.g., Lugtig, 2014). Based on research of the association between personality and the likelihood of complying with follow-up requests in surveys (Marcus & Schütz, 2005), one could, for example, expect *extraversion* and *openness to experiences* to be positively associated with participating in biomarker collections. Lastly, the amount of spare time individuals have might be related to the likelihood of biomarker participation. Related empirical evidence comes from Courvoisier et al. (2012), who showed that within-day response rates in an experience sampling study were lower during the work-day (8 am–5 pm) than after the workday (after 5 pm). In sum, various person-level and situation-level characteristics have been discussed as being related to the likelihood of biomarker participation. However, empirical studies that specifically focus on selective participation in self-collection biomarker studies that are embedded within online surveys are currently lacking.

2.3 Biomarker Collection and Panel Attrition

The third concern with the self-collection of biomarkers that are integrated in app-based panel surveys is that they potentially increase participation burden and could thus lead to higher dropout rates in the panel survey. Yet, it could also be the case that the biomarker collection makes the study participation more interesting and thus increases the commitment to continued participation. Quasi-experimental evidence on this issue comes from a study by Pashazadeh et al. (2021), who found using UKHLS data that respondents who participated in a nurse visit in wave 2 were less likely to participate in wave 3 of the main survey compared to those who were not invited to participate in the nurse visit. Interestingly, this effect was not observed in later waves. In another quasi-experimental study, Gatny et al. (2013) found that requesting and providing self-collected saliva samples triggered by a life event did not negatively affect subsequent participation in a panel study. In sum, the evidence on how participating in biomarker collection affects panel attrition is scarce and mixed. Thus, more research is needed to understand how the inclusion of biomarker self-collection in longitudinal online surveys affects panel attrition.

3 The Present Study

The present study examines the feasibility of embedding a self-administered hair collection study in an online panel survey without any face-to-face contact. The study is based on a large-scale, app-based smartphone panel survey that combined questionnaire data on well-being and health with assessments of HCC as a biomarker of chronic stress

(Hetschko et al., 2022; Lawes et al., 2023). Based on three research questions (RQ) and informed by the literature discussed above we investigate (i) participation rates in the HCC study, (ii) patterns of selective participation in the HCC study, and (iii) the association of participating in the first hair collection wave and attrition in the panel survey.

3.1 RQ1: Participation Rates in the HCC Study

The first goal of this study is to investigate participation rates in the self-administered HCC collection study. We examine both stated willingness to participate in the HCC study as well as actual participation (i.e., sending in a hair sample) in five hair collection waves.[1] Specifically, we aim to address the following research question:

RQ1: What proportion of individuals who participate in an app-based smartphone panel survey indicate their willingness to provide hair samples for cortisol measurement, and which proportion of those actually do so?

3.2 RQ2: Selective Participation in the HCC Study

The second goal is to examine individual and situational characteristics that are related to both the stated willingness to participate as well as actual participation in the HCC study conditional on stated willingness. In particular, we aim to address the following research question:

RQ2: What are individual and situational predictors of (a) stated willingness to participate in the first collection wave of the HCC study and (b) actual participation in the first collection wave conditional on willingness?

Based on the existing literature, we expect the likelihood of participation in the HCC study to increase with higher education and higher levels of perceived stress. Concerning age and gender, the existing literature shows mixed evidence (Dykema et al., 2017; Schonlau et al., 2010); thus, we refrain from deriving specific expectations in terms of these characteristics. Still, while the likelihood of participation might not differ across genders, we expect men to be underrepresented in the HCC study as they generally have shorter hair styles which do not always allow for hair sampling. Moreover, we expect that individuals who are

unemployed or expect to become unemployed are more likely to participate in HCC collection as they may have (or expect to have) more spare time compared to continuously employed individuals. Based on the same reasoning, fulltime employment and overtime work may reduce the likelihood of participating. In addition, we explore whether respondents' personality is related to participation in the HCC study.

3.3 RQ3: Participation in the First Hair Collection Wave and Attrition in the Panel Survey

The third goal of this study is to investigate whether participating in the first hair collection wave is related to panel attrition in the general smartphone survey. Specifically, we examine the relationship between participating in the first hair collection wave and panel attrition. In the process we control for the percentage of answered survey items during the first survey wave as a proxy for general motivation to participate which could influence both participation in the first hair collection wave as well as continued survey participation over many months. Thus, our third and final research question is:

RQ3: Does participating in the first hair collection wave predict long-term panel survey participation?

4 Methods

We used data from the German Job Search Panel (GJSP; Hetschko et al., 2022) to address the three research questions. The GJSP is a monthly app-based smartphone panel survey on the well-being and health of recently registered German jobseekers. The study protocol of the GJSP was approved on December 13, 2017, by the ethics committee of the Department of Education and Psychology at Freie Universität Berlin.

4.1 Sample and Data Collection

From November 2017 to May 2019, a total of 127,836 individuals aged between 18 and 60 years who registered as employed jobseekers[2] in the German unemployment insurance system were invited to participate in the GJSP over the

---

[1] Having said that, we do not analyze the reasons why certain individuals express a willingness to participate without following through later on.

[2] In Germany, employees are obliged to register as jobseekers at least three months prior to the expected day of their job loss to qualify for unemployment benefits (or within three days if they learn about the job loss later). Otherwise, a cut-off period for unemployment benefit receipt might apply.

course of up to two years. Of these identified jobseekers, 79,710 were invited because they were likely to be affected by a mass layoff,[3] whereas the remaining 48,126 individuals registered as jobseekers for a different reason (e.g., expiring contract). Until June 2018, all individuals were invited to participate in the study via a postal letter. Between July 2018 and October 2018, individuals who had provided an email address to the German Federal Employment Agency were invited via email, whereas the remaining individuals were invited via postal letter. Beginning in November 2018, individuals who had provided an email address were invited either by postal letter, email, or pre-announcement letter followed by an email for the purpose of another study (for details see Lawes, Hetschko, Sakshaug, et al., 2022). All other individuals were invited via letter during this time. To determine their eligibility for the study, individuals who were interested in participating in the GJSP could fill out an online entry survey via a downloadable survey app on a smartphone or using a web interface within 10 days after receiving the study invitation. 1540 individuals were included in the GJSP and filled out at least one questionnaire after the entry survey (see Figure S1 in the supplementary materials for a flowchart). Aggregate non-response bias for the final GJSP sample is small (for details see Hetschko et al., 2022).

The GJSP was carried out via a custom smartphone app developed by the App Research Organization (ARO), which ran on Android and iOS (Ludwigs & Erdtmann, 2019). Participants received monthly questionnaires on up to eight consecutive days via the app. The questionnaires assessed a wide range of psychological constructs and work-related variables (Hetschko et al., 2022). Crucially, individuals were informed by the study website and the recruitment flyer about the opportunity to participate in an add-on study on cortisol (for details see Material S1 in the supplementary files). It was made clear that participation in the cortisol study was not required and did not affect the eligibility to participate in the survey part of the GJSP.

4.1.1 Hair Collection Procedure

On the seventh measurement day of the first monthly survey wave (Q1), the survey app notified respondents about the role of hair cortisol as a biomarker of stress, the hair collection procedure, the additional data protection in place for the HCC study, the voluntary nature of the participation in the HCC study, and the incentive for participation in this add-on study (for details see Material S2 in the supplementary files). Respondents could then indicate via the survey app whether they were willing to participate in the HCC study. In a second step, all individuals who indicated their willingness were screened for eligibility via the app. Individuals who reported that their hair was shorter than 2 cm or that they took cortisone-based medication were deemed ineligible and excluded from the HCC study. In a third step, willing and eligible individuals were asked to provide their mailing address so that hair collection kits could be sent out via mail by the ARO. These kits contained instructions for hair removal, loops to fixate the hair strands, aluminum foil for dry and dark shipping, a prepaid return envelope, and a paper-pencil questionnaire to assess factors that may confound hair cortisol values, such as cortisone-based medication (for details see Material S3 in supplementary files).

Respondents were asked to send three hair strands (each with an overall diameter of at least 3 mm) to the research team via mail within 10 days after receiving the hair collection kit. Individuals who were willing and eligible for hair cortisol collection in Q1 received the same set of questions concerning willingness and eligibility for hair cortisol collection on a quarterly basis within the first survey year (i.e., in Q2, Q3, Q4, and Q5) regardless of whether or not they actually sent in any hair samples in Q1. Individuals who were unwilling or ineligible for the HCC study in Q1 were excluded from later hair collection waves and did not receive the HCC questionnaires anymore. Except for a few months at the end of the GJSP, hair collection kits were no longer sent out to individuals who had missed a previous HCC collection wave even if they were willing to participate and eligible. This was done in order to avoid gaps in the panel data and to maximize the number of hair samples that could be analyzed given the research budget. Toward the end of the GJPS funding period, some money was left over which we used to also analyze hair samples of individuals who had previously missed a hair sample collection wave. Respondents received a 10 euros cash incentive in the form of a bank transfer or Amazon voucher for each hair sample they sent in. Moreover, individuals could receive feedback concerning their HCC if they participated in the GJSP for at least two years.

Hair samples were analyzed by the biolaboratory Dresden Lab Service using immunoassays to obtain the cortisol concentration in the 3 cm hair segments closest to the scalp. Only very few samples did not contain sufficient material and had to be discarded (6 out of 2303 samples). Details about hair cortisol analyses are summarized by Greff et al. (2019).

---

[3] One aim of the recruitment process was to oversample people who registered as a jobseeker due to a mass layoff. To identify potential mass layoffs, we applied the thresholds for a notifiable mass layoff according to § 17(1) of the German employment protection act (*Kündigungsschutzgesetz*): >5 layoffs in plants with 21 to 59 employees, 10% layoffs in plants with 60 to 250 employees, >25 layoffs in plants with 251 to 499 employees, ≥30 layoffs in plants with 500 or more employees. In addition, we assume a mass layoff in a plant of ≤20 employees if the staff reduced by more than five people.

4.2 Analysis Strategy

To address RQ1, we computed the proportions of individuals participating and not participating due to different reasons in the five quarterly hair collection waves. To address RQ2 and RQ3, we ran several regression analyses, which will be explained below. All analyses were run in R (version 4.1.1; R Core Team, 2017) and all scripts and model results are available in the online repository of this study (https://osf.io/mt2q4/?view_only=debaee45862445a9b0b38f8710f1143c). The wordings of all items used in the analysis are presented in Material S4 of the supplementary materials.

4.2.1 Selective Participation in the HCC Study (RQ2)

We used logistic regression models to investigate whether individual and situational characteristics were related to (i) the stated willingness to participate in the first HCC collection wave and (ii) actual participation (i.e., sending in hair) conditional on willingness. Two separate dummy variables were constructed: The first variable was coded as 1 if individuals indicated that they were generally willing to participate in the first HCC collection wave, and 0 if not (*willingness*). The second variable was coded as 1 if an individual sent in their hair in the first HCC collection wave and 0 if a hair collection kit had been sent out, but no hair sample was sent back (s*ent in hair sample*). This way, the second outcome (i.e., *sent in hair sample*) is conditional on being eligible and a hair collection kit having been sent out. As independent predictor variables, we used the following measures that were all assessed in Q1.

Age, Gender, and Education. Respondents' age and gender were collected during the entry survey of the main study. Gender was assessed with the following three categories: *female, male*, and *other*. As a proxy for the level of education, a dummy variable was defined capturing whether respondents have a tertiary degree (i.e., an academic degree) or not.

Perceived Stress. Perceived stress was assessed by asking respondents to indicate how often they felt 'overburdened' and 'stressed' within the last week. Individuals could respond to both items on a five-point rating scale ranging from (1) *rarely or none of the time (less than 1 day)*, (2) *some or a little of the time (1–2 days)*, (3) *occasionally or a moderate amount of time (3–4 days)*, (4) *most or all of the time (5–7 days)*, or (5) *don't know*. *Don't know* answers were coded as missing values. The individual-specific average across the two items was transformed into percent of maximum possible scores (POMP; Cohen et al., 1999), which

range from 0 to 100 and can be interpreted in terms of percentage points.

Personality. Personality was assessed using the Big Five Inventory-SOEP (BFI-S; Schupp & Gerlitz, 2014), which measures the Big Five personality dimensions *neuroticism, extraversion, openness, agreeableness*, and *conscientiousness* with three items each. Respondents indicated on a seven-point rating scale ranging from (1) *not at all* to (7) *completely* to what extent different statements (e.g., 'I am a person who works thoroughly.') describe them. All negatively worded items were reverse-coded and the item scores were averaged across all items of a personality dimension to obtain a scale score for each subject. Analogous to the perceived stress variable, the scale scores were transformed into POMP scores.

Working Hours. Respondents were asked to indicate their contractually agreed weekly working hours. If these ranged between 0 and 15, the employment was classified as *marginally employed*. If the working hours were between 15 and 35, the employment was classified as *part-time*, and if they were more than 35 h, *full-time*. Individuals without a contractually stipulated number of working hours were classified as individuals with *flexible working hours*. Moreover, a dummy variable was created that indicated whether or not individuals worked overtime.

Employment-Related Expectations. As measures of the anticipated employment situation in the future, respondents were asked to rate on an 11-point scale ranging from 0% to 100% how likely they thought that several employment-related changes will occur within the next six months (for items see SOEP; Wagner et al., 2007). Specifically, respondents were asked how likely it is that they will *look for a new position, actually lose their job, give up their current profession and start another one, become self-employed, substantially change their working hours,* or *give up working completely*. For each of these employment-related expectations dummy variables were defined indicating whether or not the perceived likelihood of the event was 50% or greater.

Separately for the two outcomes (i.e., *willingness*, s*ent in hair sample*), we first ran bivariate regressions for each of the predictor variables. Then, the predictor variables were incorporated step-by-step in a series of multiple logistic regression models. In Model 1 self-reported perceived stress, age, gender, and education were included. Model 2 additionally included the personality variables, whereas Model 3 additionally included the working hours and employment-related expectations. Moreover, an overall model was fitted that included all predictor variables simultaneously to mutually control for the various characteristics (Model 4).

All models also contained the following control variables to account for study-specific features that might relate to participation. Most importantly, we controlled for the percentage of answered survey items during Q1 (in what follows, item response rate) as a general indicator of study compliance because we expect it to be related to the individual and situational characteristics as well as the likelihood to participate in the HCC study. Moreover, because different contact modes were used to recruit subjects for the GJSP and these contact modes are known to affect signup and participation rates (Lawes, Hetschko, Sakshaug, et al., 2022), we included dummy variables indicating with which contact mode subjects were recruited. Lastly, we included a dummy variable indicating whether a respondent was recruited within the mass layoff sample or not in order to control for the sampling procedure. All continuous variables were grand mean centered so that the intercepts of the models can be interpreted as the predicted logit for individuals with average values on the continuous covariates and values on the categorical variables that correspond to the reference categories.

The sample for all models was restricted to individuals without missing values on the covariates to ensure that the results are based on the same individuals across models. Importantly, missing values were rare so that for analyses of the outcome *willingness* only 77 (5%) individuals had to be dropped and for the outcome *sent in hair sample* 32 (4%) individuals had to be dropped. Thus, the analyses were based on $N = 1316$ individuals for the outcome *willingness* and $N = 737$ individuals for the outcome *sent in hair sample*.

4.2.2 Participation in the First Hair Collection Wave and Attrition in the Panel Survey (RQ3)

In a third step, we examined whether participating in the first hair collection wave was associated with panel attrition in the main survey. Specifically, we characterized respondents' participation in the first hair collection wave using the following categories: *not willing to participate* (reference category), *willing but not eligible, willing and eligible but no hair collection kit was sent out, hair collection kit was sent out but no hair sample was sent in, hair sample was sent in*. We used a categorical variable describing these groups in separate logistic regression models to predict two dummy outcome variables indicating whether individuals were still responding to the survey questions after 13 and 25 months. In particular, month 13 was the last collection wave of the HCC study and month 25 was the last regular wave of the GJSP. In these models we, again, controlled for the (grand mean centered) item response rate during Q1 (in %) as a general indicator of survey motivation. In this way, we aimed at ruling out that a measured commit-

ment effect of participating in the HCC study on continued survey participation originates, in fact, from a higher general motivation to comprehensively participate in surveys. Moreover, we again considered the contact mode and the sample information (i.e., mass layoff vs. no mass layoff) as study-specific control variables. Both regression models were based on all individuals who responded to the HCC questionnaire in Q1 ($N = 1393$).

5 Results

5.1 Participation Rates in the HCC Study (RQ1)

The participation outcomes for the five hair collection waves are illustrated in Table 1 and in Figures S1 (first wave) and S2 (all five waves) of the supplementary materials. Out of the 1540 individuals who started the GJSP, 1393 (90%) completed the questionnaire containing the willingness and eligibility items for the HCC study in Q1. One subject closed this questionnaire before answering the question on hair length. Of the remaining individuals, 1139 (82%) indicated their willingness to participate in the HCC collection wave. Among all willing individuals, 299 (26%) had to be excluded due to short hair ($N = 197$) or cortisone-based medication ($N = 102$), resulting in 840 eligible individuals. Of these, 71 (8%) did not provide a valid mailing address to receive the hair collection kit. Therefore, a total of 769 hair collection kits were sent out to eligible individuals during Q1. Of these, 445 (58%) were sent back with hair samples. In Q2, the second quarterly HCC collection wave, 404 hair collection kits were sent out to willing and eligible individuals, and 284 (or 70%) hair samples were received. In Q3, 219 out of 274 (or 80%) hair samples were received, in Q4, 181 out of 212 (or 85%), and in Q5, 133 out of 147 (or 90%), indicating an increasing percentage of returned hair collection kits in each subsequent wave while overall participation in the hair collection decreased.

5.2 Selective Participation in the HCC Study (RQ2)

Table 2 depicts descriptive statistics for all predictor variables for different subgroups based on participation in the first hair collection wave. It illustrates that individuals who sent in their hair samples in Q1 reported an average of 3-percentage points higher perceived stress levels and were roughly one year older compared to the overall GJSP respondent sample. Moreover, the share of females and individuals with a tertiary degree was higher in the group of individuals who sent in their hair samples in Q1 compared to all GJSP respondents. Table 3 presents the results of the lo-

Table 1

Conditional Participation Rates in the Five HCC Collection Waves (Q1–Q5)

	Q1		Q2		Q3		Q4		Q5	
	%	N	%	N	%	N	%	N	%	N
Among individuals still actively participating in the GJSP survey										
Did not respond to HCC questionnaire	9	138 [a]	3	29	1	12	4	33	2	15
Excluded because not willing or eligible in Q1	–	–	37	415	37	351	36	323	36	299
Submitted the HCC questionnaire	91	1393	60	664	62	599	60	540	62	514
Among individuals who submitted the HCC questionnaire										
Not willing to participate in hair collection	18	254	24	161	32	192	31	167	32	165
Willing to participate in hair collection	82	1139	76	503	68	407	69	373	68	349
Among individuals willing to participate in the hair collection										
Not eligible due to cortisone-based medication	9	102	2	10	5	19	5	17	8	29
Not eligible due to short hair	17	197	6	28	6	26	9	33	9	30
Eligible for hair collection	74	840	92	465	89	362	87	323	83	290
Among individuals eligible for hair collection										
No hair collection kits were mailed out[b]	9	71	13	61	24	88	34	111	49	143
Hair collection kits were mailed out	92	769	87	404	76	274	66	212	51	147
Among hair collection kits that were mailed out										
No hair sample was sent in	42	324	30	120	20	55	15	31	10	14
Hair sample was sent in	58	445	70	284	80	219	85	181	90	133

[a] Nine individuals did not fill out any survey questions in Q1
[b] For most parts of the study, hair collection kits were not sent out to individuals who had missed a previous HCC collection wave. Moreover, some individuals provided invalid address information in which case no hair collection kit was sent out

gistic regression analyses. The results show that individuals who reported higher stress levels as well as older subjects were more likely to indicate their willingness to participate in the first hair collection wave. Further, better survey participation in Q1 (indicated by higher item-level response rates) was positively associated with self-rated willingness to participate in the first HCC collection wave. Among individuals who were recruited after June 2018 and who had provided an email address during their job seeking registration, those who were contacted via some form of physical letter (i.e., letter or preannouncement letter plus email) were more likely to indicate their willingness to participate in the first HCC collection wave than individuals who were contacted via email. Moreover, individuals who were recruited until July 2018 (i.e., via letter) were more likely to indicate their willingness to participate in the first HCC collection wave than individuals who were contacted via email after June 2018. These results were consistent across the different model specifications. The other individual or situational characteristics (e.g., gender, education, personality, current employment situation) were not consistently related to self-rated willingness to participate in the first hair collection wave.

Among the willing and eligible individuals who received a hair collection kit in Q1, age and survey participation rates at Q1 were positively related to the likelihood of actually sending in a hair sample. The other individual and situational characteristics were not consistently associated with the likelihood of sending in a hair sample. Again, these results are robust across the subsetted models (i.e., bivariate, model 1–3) and the full model (model 4).

5.3 Participation in the HCC Study and Attrition in the Panel Survey (RQ3)

Table 4 presents the results of the logistic regression analysis for predicting continued panel participation in the GJSP. The results indicate that individuals who sent in their hair sample during Q1 were significantly more likely to still actively participate in the survey after one year ($b = 1.375, p < 0.001$) and after two years ($b = 1.164, p < 0.001$) compared to individuals who indicated that they were not willing to participate in the first HCC collection wave. This might reflect a generally high motivation to participate comprehensively in the study or a commitment effect originating from participating in the HCC study. Interestingly, individ-

Table 2

Descriptive Statistics at Q1 Grouped by Participation in the First HCC Collection Wave

	All GJSP respondents		Completed HCC questionaire		Willing to participapte		Eligible to participate		Sent in hair sample	
	Mean/%	S.D.	Mean/%	S.D.	Mean/%	S.D.	Mean/%	S.D.	Mean/%	S.D.
Age	38	10	39	10	39	10	39	10	40	10
Gender										
Female	52	–	52	–	53	–	63	–	66	–
Male	48	–	48	–	47	–	37	–	34	–
Other gender	0	–	0	–	0	–	0	–	0	–
Tertiary degree	46	–	47	–	46	–	49	–	52	–
Married	43	–	44	–	45	–	44	–	46	–
Household income in Euro	2830	1673	2843	1690	2870	1708	2871	1728	2874	1628
Perceived stress (*POMP score*)	34	30	35	30	36	30	36	30	37	31
Personality (*POMP score*)										
Neuroticism	50	21	50	21	50	21	50	21	50	21
Openness	66	18	66	18	66	18	66	18	66	19
Extraversion	63	20	63	20	63	20	63	20	63	20
Agreeableness	75	16	75	16	75	15	75	15	76	15
Conscientiousness	80	14	80	14	80	14	80	14	80	14
Employment situation										
Full-time employment	60	–	59	–	58	–	56	–	54	–
Marginal employment	1	–	1	–	1	–	1	–	1	–
Part-time employment	35	–	35	–	36	–	38	–	40	–
Flexible working hours	5	–	5	–	5	–	4	–	4	–
Works overtime	24	–	24	–	24	–	24	–	26	–
Within six months, respondent expects to										
...look for a new position	67	–	68	–	68	–	66	–	68	–
...lose job	55	–	55	–	56	–	55	–	55	–
...start a new profession	25	–	25	–	25	–	26	–	26	–

Table 2

(*Continued*)

	All GJSP respondents		Completed HCC questionaire		Willing to participate		Eligible to participate		Sent in hair sample	
	Mean/%	S.D.	Mean/%	S.D.	Mean/%	S.D.	Mean/%	S.D.	Mean/%	S.D.
... become self-employed	6	–	6	–	6	–	7	–	6	–
... change working hours	13	–	14	–	14	–	16	–	17	–
... give up working entirely	7	–	7	–	7	–	7	–	6	–
% of answered items in Q1	89	16	92	9	93	9	93	9	94	6
Contact mode for recruitment										
Letter (12/2017-06/2018)[a]	14	–	14	–	15	–	15	–	16	–
Letter (07/2018-05/2019)[b]	14	–	14	–	13	–	13	–	13	–
Letter (11/2018-05/2019)[c]	15	–	15	–	17	–	16	–	16	–
Email	33	–	32	–	30	–	31	–	32	–
Pre-Announcement	24	–	25	–	26	–	25	–	23	–
Mass layoff sample	62	–	62	–	62	–	62	–	65	–
N	1540	–	1393	–	1139	–	840	–	445	–

For categorical variables proportions are depicted. For continuous variables, the mean and the standard deviation are presented.
[a] Letter from Dec. 2017 to June 2018, unclear if email address was provided, [b] Letter from July 2018 to May 2019, no email address provided, [c] Letter from Nov. 2018 to May 2019, email address provided

Table 3

Logistic Regression Coefficients (Log-Odds Ratios and Standard Errors) of Willingness to Participate and Actual Participation in the First HCC Collection Wave

	Willingness					Sent in hair sample				
	Bivariate	Model 1	Model 2	Model 3	Model 4	Bivariate	Model 1	Model 2	Model 3	Model 4
Age	0.022*** (0.007)	0.024** (0.007)	–	–	0.026** (0.008)	0.025*** (0.008)	0.031** (0.008)	–	–	0.032** (0.008)
Gender (ref: female)										
Male	–0.176 (0.143)	–0.186 (0.149)	–	–	–0.196 (0.159)	–0.304 (0.155)	–0.276 (0.161)	–	–	–0.288 (0.173)
Other gender	13.063 (509.652)	13.448 (486.576)	–	–	13.019 (492.969)	13.27 (535.411)	13.219 (535.411)	–	–	12.993 (535.411)
Tertiary degree (ref.: no tertiary degree)	–0.032 (0.143)	–0.103 (0.153)	–	–	–0.140 (0.159)	0.272 (0.149)	0.303 (0.162)	–	–	0.293 (0.167)
Perceived stress (POMP score)	0.005* (0.002)	0.006* (0.003)	–	–	0.006* (0.003)	0.004 (0.003)	0.004 (0.003)	–	–	0.005 (0.003)
Personality (POMP scores)										
Neuroticism	0.001 (0.003)	–	0.001 (0.004)	–	–0.001 (0.004)	–0.002 (0.004)	–	–0.001 (0.004)	–	–0.003 (0.004)
Openness	0.001 (0.004)	–	0.003 (0.004)	–	0.0004 (0.004)	–0.002 (0.004)	–	–0.002 (0.004)	–	–0.003 (0.005)
Extraversion	–0.001 (0.004)	–	–0.0002 (0.004)	–	–0.0001 (0.004)	–0.001 (0.004)	–	0.0003 (0.004)	–	0.0003 (0.004)
Agreeableness	0 (0.005)	–	0.001 (0.005)	–	0.001 (0.005)	0.002 (0.005)	–	0.002 (0.005)	–	0.002 (0.005)
Conscientiousness	–0.008 (0.005)	–	–0.008 (0.006)	–	–0.012* (0.006)	–0.001 (0.006)	–	–0.0004 (0.006)	–	–0.005 (0.006)
Employment situation (ref: full-time employment)										
Marginal employment	1.102 (1.044)	–	–	1.202 (1.076)	1.163 (1.134)	–0.188 (0.64)	–	–	–0.016 (0.684)	–0.272 (0.713)
Part-time employment	0.271 (0.154)	–	–	0.257 (0.160)	0.193 (0.166)	0.256 (0.156)	–	–	0.227 (0.161)	0.122 (0.169)
Flexible working hours	0.675 (0.443)	–	–	0.731 (0.457)	0.708 (0.464)	0.448 (0.424)	–	–	0.442 (0.435)	0.404 (0.444)
Works overtime (ref.: does not work overtime)	0.104 (0.172)	–	–	0.087 (0.178)	0.070 (0.181)	0.209 (0.176)	–	–	0.201 (0.182)	0.258 (0.188)

Table 3

(Continued)

	Willingness					Sent in hair sample				
	Bivariate	Model 1	Model 2	Model 3	Model 4	Bivariate	Model 1	Model 2	Model 3	Model 4
Within six months, respondent expects to (ref.: does not expect)										
... look for a new position	0.071 (0.151)	—	—	-0.071 (0.186)	-0.025 (0.189)	0.256 (0.157)	—	—	0.397* (0.194)	0.375 (0.200)
... lose job	0.107 (0.143)	—	—	0.137 (0.173)	0.093 (0.175)	-0.013 (0.15)	—	—	-0.176 (0.181)	-0.226 (0.186)
... start a new profession	0.18 (0.172)	—	—	0.174 (0.186)	0.197 (0.189)	0.002 (0.169)	—	—	-0.055 (0.185)	-0.009 (0.190)
... become self-employed	0.804* (0.404)	—	—	0.797 (0.417)	0.766 (0.421)	-0.182 (0.302)	—	—	-0.159 (0.313)	-0.209 (0.327)
... change working hours	0.232 (0.223)	—	—	0.059 (0.233)	0.037 (0.238)	0.144 (0.206)	—	—	0.078 (0.218)	0.067 (0.223)
... give up working entirely	-0.458 (0.248)	—	—	-0.581* (0.258)	-0.525* (0.263)	-0.238 (0.288)	—	—	-0.165 (0.299)	-0.074 (0.306)
% of answered items in Q1	0.015* (0.007)	0.019** (0.007)	0.016* (0.007)	0.018* (0.007)	0.021** (0.007)	0.04*** (0.013)	0.044** (0.013)	0.040** (0.013)	0.038** (0.013)	0.042** (0.013)
Contact mode for recruitment (ref: email)										
Letter from July 2018 to May 2019, no email address provided	-0.053 (0.206)	-0.104 (0.213)	-0.030 (0.208)	-0.059 (0.211)	-0.124 (0.217)	0.097 (0.246)	0.180 (0.255)	0.121 (0.250)	0.160 (0.253)	0.181 (0.262)
Letter from Dec. 2017 to June 2018, unclear if email address was provided	0.673*** (0.252)	0.742** (0.267)	0.768** (0.263)	0.765** (0.265)	0.749** (0.270)	0.05 (0.239)	0.075 (0.263)	-0.036 (0.258)	-0.026 (0.259)	0.038 (0.267)
Pre-announcement letter and email	0.524*** (0.192)	0.541** (0.195)	0.530** (0.193)	0.508** (0.194)	0.536** (0.197)	-0.094 (0.198)	-0.105 (0.205)	-0.151 (0.202)	-0.107 (0.204)	-0.089 (0.210)
Letter from Nov. 2018 to May 2019, email address provided	0.844*** (0.25)	0.847** (0.252)	0.852** (0.251)	0.817** (0.253)	0.793** (0.255)	0.013 (0.23)	0.009 (0.237)	-0.025 (0.235)	-0.025 (0.236)	-0.035 (0.242)
Mass layoff sample (ref.: not mass layoff sample)	-0.023 (0.147)	-0.103 (0.157)	-0.137 (0.155)	-0.177 (0.157)	-0.145 (0.160)	0.273 (0.153)	0.259 (0.167)	0.284 (0.164)	0.270 (0.165)	0.240 (0.170)
Constant	—	0.346 (0.428)	0.414 (0.402)	0.135 (0.430)	0.162 (0.458)	—	-2.416** (0.756)	-2.078** (0.725)	-2.296** (0.733)	-2.486** (0.770)
Observations	1316	1316	1316	1316	1316	737	737	737	737	737

*p < 0.05, **p < 0.01

Model 1: model including variables with a clear hypothesis; Model 2: model including personality variables; Model 3: model with employment-related variables; Model 4: full model; Models 1–4 contain study-specific control variables

Table 4

Logistic Regression Coefficients (Log-Odds Ratios and Standard Errors) of Long-Term Panel Participation

	Participated in survey after 13 months		Participated in survey after 25 months	
	Coef.	Std. Err.	Coef.	Std. Err.
HCC participation outcomes in Q1 (ref.: not willing to participate)				
Willing but not eligible	0.586**	0.180	0.734**	0.201
Willing and eligible but no hair collection kit was mailed out[a]	−0.331	0.281	−0.244	0.356
No hair sample was sent in	0.139	0.174	0.106	0.208
Hair sample was sent in	1.375**	0.176	1.164**	0.186
% of answered items in Q1 (centered)	0.032**	0.008	0.025**	0.010
Contact mode for recruitment (ref.: email)				
Letter from July 2018 to May 2019, no email address provided	0.349	0.187	0.302	0.190
Letter from Dec. 2017 to June 2018, unclear if email address was provided	−0.849**	0.195	−0.150	0.207
Pre-announcement letter and email	0.467**	0.159	0.273	0.159
Letter from Nov. 2018 to May 2019, email address provided	0.341	0.183	0.079	0.186
Mass layoff sample (ref.: not mass layoff sample)	0.181	0.127	0.013	0.129
Constant	−2.056**	0.477	−2.885**	0.566
Observations	1393		1393	

$^{*} p < 0.05, ^{**} p < 0.01$

[a] For most parts of the study, hair collection kits were not sent out to individuals who had missed a previous HCC collection wave. Moreover, some individuals provided invalid addresses so that no hair collection kit was sent out

uals who indicated their willingness to participate in the first hair collection wave but who had to be excluded from it due to a short hair style or the intake of cortisone-based medication were also more likely to still actively participate in the survey after 13 months ($b = 0.586$, $p = 0.001$) and 25 months ($b = 0.734$, $p < 0.001$) compared to individuals who were unwilling to participate in the first HCC collection wave. For these individuals, a commitment effect originating from active participation in the HCC study can be ruled out. However, when comparing the two groups of willing individuals with each other, those who sent in their hair sample during Q1 were still significantly more likely to actively participate in the survey after one year ($b = 0.789$, $p < 0.001$) and after two years ($b = 0.430$, $p = 0.006$) than those who indicated their willingness to participate but had to be excluded due to not meeting the eligibility criteria (post-estimation test not shown in Table 4). This holds true even though the item response rate at Q1 was controlled for.[4] Lastly, individuals who were recruited into the GJSP during the earlier months of the GJSP (i.e., via letter) were less likely ($b = –0.849$, $p < 0.001$), whereas individuals contacted via pre-announcement ($b = 0.467$, $p = 0.003$) were more likely, to still participate in the survey after one year compared to individuals contacted via email.

6   Discussion

This is the first study that examined the feasibility of incorporating a self-collection biomarker study into an online panel survey without any face-to-face interaction with its respondents. Specifically, we investigated participation rates and patterns of selective participation in a longitudinal hair cortisol concentration (HCC) study that was embedded within an app-based smartphone panel survey on the well-being and health of German jobseekers. Further, we examined whether participating in the first hair collection wave was associated with panel attrition in the survey.

6.1   Participation Rates in the HCC Study (RQ1)

The vast majority (82%) of respondents indicated that they were willing to participate in the first hair collection wave. After excluding individuals with short hair (17%) and individuals who took cortisone-based drugs (9%), about 74% percent of individuals who indicated their willingness to

participate in the HCC study also met the eligibility criteria. As expected, these eligibility-based exclusions resulted in some selectivity. Specifically, the share of females in the samples increased from 52% to 62%. This finding is in line with previous studies and shows an important drawback of hair sampling studies, namely that individuals with short hair styles cannot participate.

Another important finding of this study was that among all willing and eligible individuals who received a hair collection kit only 58% followed through and eventually sent in their hair samples in the first hair collection wave. Reassuringly, the return rates in the subsequent quarterly HCC collection waves were higher (i.e., ranging from 70–90%) even though overall participation decreased over time. This observed increase in the return rates likely originated from different processes. First, it is likely that many participants knew better what to expect from the hair collection procedure after having participated in it at least once before, so that they were more likely to only indicate their willingness to participate in the later hair collection waves if they really intended to send in their hair samples. Empirical evidence for this explanation comes from the lower proportions of individuals who expressed their willingness to participate in later hair collection waves. A second reason for the higher return rates in later hair collection waves is that hair collection kits were generally only mailed out to individuals who had not missed a past hair collection wave, so that only previously compliant individuals could participate in the later waves of the HCC study.

6.2   Selective Participation in the HCC Study (RQ2)

The stated willingness to participate in the first hair collection wave was selective to a limited extent. Older individuals as well as individuals with higher self-reported levels of perceived stress were more likely to express their willingness to participate in the first hair collection wave. These results are in line with our expectations as well as existing research by Sakshaug et al. (2010) in that individuals with pronounced self-rated levels of the outcome variable measured by the biomarker (here: stress) were more likely to be willing to participate in the biomarker collection. The positive effect of age on the stated willingness to participate in the HCC study is further consistent with findings reported by Schonlau et al. (2010). Importantly, these selection effects were comparably small in our study, which becomes apparent when the logit coefficients are transformed into odds-ratios. In particular, an individual who is 10 years older than a person with otherwise identical values on the covariates is predicted to be only 1.3 times more likely to indicate their willingness to participate in the HCC collection. Analogously, a person who indicates a one-cate-

---

[4] Consistent with our notion that the survey item response rate during Q1 is a suitable proxy measure of general study compliance, the regressions indicated that it was positively related to the likelihood of continuously participating in the survey after 13 months ($b = 0.032$, $p < 0.001$) and 25 months ($b = 0.025$, $p = 0.009$).

gory greater stress level on a four-point rating scale (i.e., 25 POMP scores) is predicted to be only 1.2 times more likely to express a willingness to participate in the HCC study than a person who has identical values on the other covariates. Crucially, against our expectations none of the other individual-level or situational characteristics examined in this study (i.e., personality, employment situation, time availability) consistently predicted the stated willingness to participate in the HCC study. When it comes to actual participation in the HCC study (and not merely hypothetical willingness to participate), only age was weakly and positively related to the likelihood of actually sending in hair samples after having received a hair collection kit. Thus, overall, these results are reassuring and indicate only limited selectivity in terms of both the general willingness to participate in the HCC study as well as actual participation after having received the hair collection kit. Put differently, while some individual characteristics are predictive of participation in the GJSP generally (Hetschko et al., 2022), participation in the HCC add-on study is hardly selective above and beyond that.

6.3 Participation in the HCC Study and Attrition in the Panel Survey (RQ3)

Our study suggests that participating in the first hair collection wave was negatively associated with panel attrition. Specifically, individuals who sent in their hair samples in the first collection wave were roughly four times more likely to still participate in the panel study after one year than those who indicated that they were generally not willing to participate in the first hair collection wave (but had identical values on the other covariates). Importantly, this large positive effect was present even though an indicator of general study compliance in the first survey wave were controlled for. Moreover, individuals who indicated that they were willing to partake in the first hair collection wave but were excluded due to short hair or the intake of cortisone-based medication were also more likely to remain active in the smartphone survey compared to individuals who were not willing to participate in the first hair collection wave. Crucially, however, individuals who actually participated in the first hair collection wave were more likely to still participate in the panel survey after one and two years than willing individuals who had to be excluded. Thus, whereas the willingness to participate in the first hair collection wave was already related to survey compliance, actually participating in it seems to be an even stronger predictor of long-term survey participation. In contrast, individuals who did not send in hair after having received the hair collection kit were just as likely to stop participating in the survey after one and two years as individuals who were not will-

ing to participate in the first collection wave making these behaviors important warning signs of future dropout. Thus, the concern that participation burden accumulates among respondents who participate in a biomarker add-on study (Pashazadeh et al., 2021) may be unwarranted. Rather, our study suggests the opposite, namely that participating in the biomarker study might actually increase commitment to a survey. While our study provides initial evidence for this effect, more studies that are specifically designed to investigate how embedding biomarker studies in panel surveys affect long-term panel participation are needed.

6.4 Implications for Applied Researchers

The study underlines that it is possible to recruit subjects for a HCC study via a smartphone app without any face-to-face interaction. This valuable feature makes the recruitment process and the biomarker collection itself highly efficient and is likely to generalize to other biomarkers that can be self-administered (e.g., nasal swaps, see Akmatov & Pessler, 2011; saliva, see Fernandes et al., 2013). Moreover, survey respondents' general willingness to participate in the HCC study appeared to be high and hardly selective. The strongest selectivity was induced by the inevitable exclusion of individuals with short hair styles, which lead to an underrepresentation of men. Thus, before conducting biomarker studies researchers should explore the specific requirements of collecting different biomarkers to minimize exclusion-based selection effects. In terms of hair sample collection, for example, researchers could inform the participants about the planned hair collection dates in advance so that participants can plan their haircuts accordingly. Further, applied researchers might want to analytically correct for selection effects (e.g., through weighting) when analyzing the data of hair collection studies.

Another central finding of this study is the large incongruence between respondents' purported willingness to participate in the HCC study and their actual participation (i.e., sending in hair). Even though this issue is in line with existing studies (e.g., Bosnjak et al., 2005; Struminskaya et al., 2021), it seems particularly severe in our study. We would have expected roughly twice as many hair samples to be collected during the first collection wave based on the stated willingness to participate in the HCC study. Somewhat reassuringly, the return rates increased in subsequent collection waves (while overall participation rates decreased). Thus, researchers considering to explore the feasibility of biomarker collection simply by asking survey respondents about hypothetical willingness should be cautious and keep in mind that a significant share of respondents who ostensibly indicate that they are willing to participate will likely not actually send in biomarker material. Therefore, applied

researchers should motivate eligible individuals to follow through on their willingness to participate. This challenge seems especially relevant for online biomarker studies that lack any face-to-face interaction with subjects. To maximize the participation researchers could also send the biomarker collection kits to all potential participants, instead of asking for willingness first. This approach has the advantage that participants have access to all information about the collection procedure before they have to decide whether they are willing to participate in the biomarker collection. However, this approach will likely also increase costs (more kits would need to be shipped) and could raise ethical concerns.

Finally, the present study indicates that participating in the first hair collection wave was positively related to long-term panel participation even when controlling for general study compliance. Thus, integrating biomarker collections into panel surveys might not only allow for collecting informative data for research but also make the survey more interesting and attractive for long-term participation. On the contrary, not taking part in the biomarker collection might be an important warning sign of future attrition, allowing researchers to flag individuals who may require further motivation to continue participating in the survey (e.g., through additional incentives).

6.5   Limitations and Future Research

The present study was embedded in a large smartphone panel survey of a diverse population and can thus present insights into the feasibility and participation patterns of conducting an online HCC study without any face-to-face interaction with its respondents. However, the survey was not directly designed to examine why individuals (did not) participate in the hair collections. Studies with detailed questionnaires or in-depth qualitative interviews investigating the motivation and rationales of individuals are thus needed to further understand the underlying processes. In particular, it seems important to explain why numerous people were willing to participate in the HCC study at first, but then did not follow through with sending in samples. Additionally, further research is needed to investigate ways to effectively design the materials to inform the participants about the hair collection procedure, simplify the collection as much as possible for them as well as provide support if they encounter any problems.

Moreover, it would be interesting to follow up on the finding that participating in the first hair collection wave was positively associated with long-term panel participation. Specifically, it would be valuable to directly examine whether participating in an add-on biomarker collection study actually increases overall commitment to the panel survey or if generally highly committed individuals are just

more likely to participate in the hair collections. Causal identification in this manner might be achieved through an experimental setting where only a random subsample is invited to participate in the biomarker study (see Pashazadeh et al., 2021). Lastly, an important avenue for future research is to examine ways to collect other biomarkers (e.g., saliva, blood spots/pressure) anonymously through self-administered surveys. Such insights would open up new possibilities to effectively collect biomarkers in large-scale panel surveys, which may then lead to a differentiated measurement of health and a more holistic understanding of health-related correlates.

References

Akmatov, M. K., & Pessler, F. (2011). Self-collected nasal swabs to detect infection and colonization: A useful tool for population-based epidemiological studies? *International Journal of Infectious Diseases*, *15*(9), e589–e593.   https://doi.org/10.1016/j.ijid.2011.04.009.

Avendano, M., Scherpenzeel, A., & Mackenbach, J. P. (2011). Can biomarkers be collected in an internet survey? A pilot study in the LISS panel. In M. Das, P. Ester & L. Kaczmirek (Eds.), *Social Research and the Internet* (pp. 371–412). Taylor & Francis. https://doi.org/10.4324/9780203844922-15.

Börsch-Supan, A., Hank, K., & Jürges, H. (2005). A new comprehensive and international view on ageing: Introducing the "Survey of Health, Ageing and Retirement in Europe". *European Journal of Ageing*, *2*(4), 245–253. https://doi.org/10.1007/s10433-005-0014-9.

Bosnjak, M., Tuten, T. L., & Wittmann, W. W. (2005). Unit (non)response in web-based access panel surveys: An extended planned-behavior approach. *Psychology and Marketing*, *22*(6), 489–505. https://doi.org/10.1002/mar.20070.

Boyle, J., Berman, L., Dayton, J., Iachan, R., Jans, M., & ZuWallack, R. (2021). Physical measures and biomarker collection in health surveys: Propensity to participate. *Research in Social and Administrative*

Pharmacy, *17*(5), 921–929. https://doi.org/10.1016/j.sapharm.2020.07.025.

Chrousos, G. P. (2009). Stress and disorders of the stress system. *Nature Reviews Endocrinology*, *5*(7), 374–381. https://doi.org/10.1038/nrendo.2009.106.

Cohen, P., Cohen, J., Aiken, L. S., & West, S. G. (1999). The problem of units and the circumstance for POMP. *Multivariate Behavioral Research*, *34*(3), 315–346. https://doi.org/10.1207/S15327906MBR3403_2.

Courvoisier, D. S., Eid, M., & Lischetzke, T. (2012). Compliance to a cell phone-based ecological momentary assessment study: The effect of time and personality characteristics. *Psychological Assessment*, *24*(3), 713–720. https://doi.org/10.1037/a0026733.

Dykema, J., DiLoreto, K., Croes, K. D., Garbarski, D., & Beach, J. (2017). Factors associated with participation in the collection of saliva samples by mail in a survey of older adults. *Public Opinion Quarterly*, *81*(1), 57–85. https://doi.org/10.1093/poq/nfw045.

Etter, J. F., & Bullen, C. (2011). Saliva cotinine levels in users of electronic cigarettes. *European Respiratory Journal*, *38*(5), 1219–1220. https://doi.org/10.1183/09031936.00066011.

Fernandes, A., Skinner, M. L., Woelfel, T., Carpenter, T., & Haggerty, K. P. (2013). Implementing self-collection of biological specimens with a diverse sample. *Field Methods*, *25*(1), 58–73. https://doi.org/10.1177/1525822X12453526.

Ford, J. L., Boch, S. J., & McCarthy, D. O. (2016). Feasibility of hair collection for cortisol measurement in population research on adolescent health. *Nursing Research*, *65*(3), 249–255. https://doi.org/10.1097/NNR.0000000000000154.

Gatny, H. H., Couper, M. P., & Axinn, W. G. (2013). New strategies for biosample collection in population-based social research. *Social Science Research*, *42*(5), 1402–1409. https://doi.org/10.1016/j.ssresearch.2013.03.004.

Greff, M. J. E., Levine, J. M., Abuzgaia, A. M., Elzagallaai, A. A., Rieder, M. J., & van Uum, S. H. M. (2019). Hair cortisol analysis: An update on methodological considerations and clinical applications. *Clinical Biochemistry*, *63*, 1–9. https://doi.org/10.1016/j.clinbiochem.2018.09.010.

Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience sampling method: Measuring the quality of everyday life*. SAGE.

Hetschko, C., Schmidtke, J., Eid, M., Lawes, M., Schöb, R., & Stephan, G. (2022). The German Job Search Panel. *OSF Preprint.*. https://doi.org/10.31219/osf.io/7jazr.

Heuser, I., Lammers, C. H., Toussaint, Thijssen, Finch, Swaab, De Kloet, Kessler, Saletu, De Wied, & Ni-jhuis (2003). Stress and the brain. *Neurobiology of Aging*, 24. https://doi.org/10.1016/S0197-4580(03)00048-4.

Jäckle, A., Burton, J., Couper, M. P., & Lessof, C. (2019). Participation in a mobile app survey to collect expenditure data as part of a large-scale probability household panel: coverage and participation rates and biases. *Survey Research Methods*, *13*(1), 23–44. https://doi.org/10.18148/srm/2019.v13i1.7297.

Juster, F. T., & Suzman, R. (1995). An overview of the health and retirement study. *The Journal of Human Resources*, *30*, S7–S56. https://doi.org/10.2307/146277.

Keusch, F., Struminskaya, B., Antoun, C., Couper, M. P., & Kreuter, F. (2019). Willingness to participate in passive mobile data collection. *Public Opinion Quarterly*, *83*(S1), 210–235. https://doi.org/10.1093/poq/nfz007.

Kirschbaum, C., Tietze, A., Skoluda, N., & Dettenborn, L. (2009). Hair as a retrospective calendar of cortisol production–Increased cortisol incorporation into hair in the third trimester of pregnancy. *Psychoneuroendocrinology*, *34*(1), 32–37. https://doi.org/10.1016/j.psyneuen.2008.08.024.

Lawes, M., Hetschko, C., Sakshaug, J. W., & Grießemer, S. (2022). Contact modes and participation in app-based smartphone surveys: Evidence from a large-scale experiment. *Social Science Computer Review*, *40*(5), 1076–1092. https://doi.org/10.1177/0894439321993832.

Lawes, M., Hetschko, C., Schöb, R., Stephan, G., & Eid, M. (2022). Unemployment and hair cortisol as a biomarker of chronic stress. *Scientific Reports*, 12(21573), 1–10. https://doi.org/10.1038/s41598-022-25775-1.

Lawes, M., Hetschko, C., Schöb, R., Stephan, G., & Eid, M. (2023). The impact of unemployment on cognitive, affective, and eudaimonic well-being facets: Investigating immediate effects and short-term adaptation. *Journal of Personality and Social Psychology*, *124*(3), 659–681. https://doi.org/10.1037/pspp0000417.

Ludwigs, K., & Erdtmann, S. (2019). The Happiness Analyzer—developing a new technique for measuring subjective well-being. *International Journal of Community Well-Being*, *1*(2), 101–114. https://doi.org/10.1007/s42413-018-0011-3.

Lugtig, P. (2014). Panel attrition: Separating tayers, fast attriters, gradual attriters, and lurkers. *Sociological Methods and Research*, *43*(4), 699–723. https://doi.org/10.1177/0049124113520305.

Marcus, B., & Schütz, A. (2005). Who are the people reluctant to participate in research? Personality correlates

of four different types of nonresponse as inferred from self- and observer ratings. *Journal of Personality*, *73*(4), 959–984. https://doi.org/10.1111/j.1467-6494.2005.00335.x.

McGeeney, K., & Weisel, R. (2015). App vs. Web for surveys of smartphone users: Experimenting with mobile apps for signal-contingent experience sampling method surveys. In *Pew Research Center report* (pp. 77–87). https://www.pewresearch.org/methods/2015/04/01/app-vs-web-for-surveys-of-smartphone-users/.

Miller, G. (2012). The smartphone psychology manifesto. *Perspectives on Psychological Science*, *7*(3), 221–237. https://doi.org/10.1177/1745691612441215.

OECD (2013). *OECD guidelines on measuring subjective well-being*. OECD Publishing. https://doi.org/10.1787/9789264191655-en.

Pashazadeh, F., Cernat, A., & Sakshaug, J. (2021). The effects of biological data collection in longitudinal surveys on subsequent wave nonresponse. In P. Lynn (Ed.), *Advances in Longitudinal Survey Methodology* (pp. 100–121). John Wiley & Sons.

Penz, M., Wekenborg, M. K., Pieper, L., Beesdo-Baum, K., Walther, A., Miller, R., Stalder, T., & Kirschbaum, C. (2018). The Dresden Burnout Study: Protocol of a prospective cohort study for the bio-psychological investigation of burnout. *International Journal of Methods in Psychiatric Research*, *27*(2), 1–11. https://doi.org/10.1002/mpr.1613.

Perry, B., Herrington, W., Goldsack, J. C., Grandinetti, C. A., Vasisht, K. P., Landray, M. J., Bataille, L., DiCicco, R. A., Bradley, C., Narayan, A., Papadopoulos, E. J., Sheth, N., Skodacek, K., Stem, K., Strong, T. V., Walton, M. K., & Corneli, A. (2018). Use of mobile devices to measure outcomes in clinical research, 2010–2016: A systematic literature review. *Digital Biomarkers*, *2*(1), 11–30. https://doi.org/10.1159/000486347.

R Core Team (2017). *R: A language and environment for statistical computing [Computer software]*. R Foundation for Statistical Computing. https://www.r-project.org/

Revilla, M., Couper, M. P., & Ochoa, C. (2019). Willingness of online panelists to perform additional tasks. *Methods, Data, Analyses*, *13*(2), 223–251. https://doi.org/10.12758/mda.2018.01.

Russell, E., Koren, G., Rieder, M., Van Uum, S., Uum, S. V., & Van Uum, S. (2012). Hair cortisol as a biological marker of chronic stress: Current status, future directions and unanswered questions. *Psychoneuroendocrinology*, *37*(5), 589–601. https://doi.org/10.1016/j.psyneuen.2011.09.009.

Sakshaug, J. W., Couper, M. P., & Ofstedal, M. B. (2010). Characteristics of physical measurement consent in a population-based survey of older adults. *Medical Care*, *48*(1), 64–71. https://doi.org/10.1097/MLR.0b013e3181adcbd3.

Sakshaug, J. W., Ofstedal, M. B., Guyer, H., & Beebe, T. J. (2015). The collection of biospecimens in health surveys. In T. Johnson (Ed.), *Handbook of Health Survey Methods* (pp. 383–419). John Wiley & Sons. https://doi.org/10.1002/9781118594629.ch15.

Schaafsma, F. G., Hulsegge, G., de Jong, M. A., Overvliet, J., van Rossum, E. F. C., & Nieuwenhuijsen, K. (2021). The potential of using hair cortisol to measure chronic stress in occupational healthcare; A scoping review. *Journal of Occupational Health*, *63*(1), e12189. https://doi.org/10.1002/1348-9585.12189.

Schonlau, M., Reuter, M., Schupp, J., Montag, C., Weber, B., Dohmen, T., Siegel, N. A., Sunde, U., Wagner, G. G., & Falk, A. (2010). Collecting genetic samples in population wide (panel) surveys: Feasibility, nonresponse and selectivity. *Survey Research Methods*, *4*(2), 121–126. https://doi.org/10.18148/srm/2010.v4i2.3959.

Schupp, J., & Gerlitz, J. (2014). *Big Five Inventory-SOEP (BFI-S)*. Zusammenstellung Sozialwissenschaftlicher Items Und Skalen (ZIS). https://doi.org/10.6102/zis54.

Stalder, T., & Kirschbaum, C. (2012). Analysis of cortisol in hair—State of the art and future directions. *Brain, Behavior, and Immunity*, *26*(7), 1019–1029. https://doi.org/10.1016/j.bbi.2012.02.002.

Struminskaya, B., Lugtig, P., Toepoel, V., Schouten, B., Giesen, D., & Dolmans, R. (2021). Sharing data collected with smartphone sensors. *Public Opinion Quarterly*, *85*, 423–462. https://doi.org/10.1093/poq/nfab025.

Tafet, G. E., & Nemeroff, C. B. (2016). The links between stress and depression: Psychoneuroendocrinological, genetic, and environmental interactions. *Journal of Neuropsychiatry and Clinical Neurosciences*, *28*(2), 77–88. https://doi.org/10.1176/appi.neuropsych.15030053.

University of Essex, Institute for Social and Economic Research (2022). *Understanding Society: Waves 1–11, 2009–2020 and harmonised BHPS: Waves 1–18, 1991–2009. [Data collection]* (16th edn.). UK Data Service. https://doi.org/10.5255/UKDA-SN-6614-17.

Wagner, G. G., Frick, J. R., & Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP)—Scope, evolution and enhancements. *Schmollers Jahrbuch*, *127*, 139–169.

Weinstein, M., Vaupel, J. W., & Wachter, K. W. (2008). *Biosocial Surveys*. National Academies Press (US).