Survey Research Methods (2010)
Vol.4, No.2, pp. 121-126
ISSN 1864-3361
http://www.surveymethods.org

Collecting Genetic Samples in Population Wide (Panel) Surveys: Feasibility, Nonresponse and Selectivity

Matthias Schonlau[1,2,3], Martin Reuter[4,11], Jürgen Schupp[1,5,6], Christian Montag[4], Bernd Weber[1,4,11], Thomas Dohmen[1,6,8], Nico A. Siegel[7], Uwe Sunde[1,6,9], Gert G. Wagner[1,2,6,10], and Armin Falk[1,4,6,11]

[1]DIW Berlin (German Institute for Economic Research), Germany, [2]Max Planck Institute for Human Development, Berlin, Germany, [3]RAND Corporation, Pittsburgh, USA, [4]University of Bonn, Germany, [5]Free University of Berlin, Germany, [6]IZA, Bonn, Germany, [7]TNS Infratest Sozialforschung, Munich, Germany, [8]Maastricht University, Netherlands, [9]University of St. Gallen, Switzerland, [10]Berlin University of Technology, Germany, [11]Center for Economics & Neuroscience, University of Bonn

Collecting biomarkers as part of general purpose surveys offers scientists – and social scientists in particular – the ability to study biosocial phenomena, e.g. the relation between genes and human behavior. We explore the feasibility of collecting buccal cells for genetic analyses with normal interviewers as part of a pretest for the German Socio-economic Panel Study (SOEP) using a probability sample. We introduce a new non-invasive technique for collecting cell material for genetic analyses and test its quality. We found no technical difficulties in collecting the genetic material and almost all samples collected could be analyzed. However, one third of interviewers reported it was hard to convince panel members to participate. The "biomarker wave" showed no reduction in response rate compared to the previous wave that included no biomarkers and the sample exhibited very little selectivity. We conclude that collecting cell material for genetic analyses with normal interviewers is feasible with no apparent same-wave attrition, though so far we cannot rule out attrition in subsequent waves.
Keywords: biomarkers, genetic material, surveys, panel studies, non-response

1 Introduction

Scientists and social scientists in particular are increasingly interested in collecting biomarkers such as grip strength or blood as part of surveys because biomarkers offer more detailed and more objective measure of, for example, respondents' health than self assessed measures can provide. Genetic material can be extracted from some biological samples (e.g. blood, buccal cells). This in turn enables scientists to study the relationship between genes and human behavior. The hopes are high: "Biomarkers on social surveys may well reveal more about subjects' predispositions and their ancestry than do their verbal responses on which social scientists have historically depended" (Butz and Torrey, 2006). However, collecting biomarkers may also have unintended and potentially serious consequences. Because collecting biomarkers is a different and potentially more invasive request than responding to a questionnaire and because of respondents' concerns about privacy and data protection, collecting biomarkers may cause selective attrition and thus lead to biased inference.

The generalizability of biosocial research hinges on having a probability sample. A selective sample may offer limited variability in the behavioral domain of interest. A probability sample is desirable in all surveys but may be of particular importance if the gene-environment interaction of interest is related to decision making (Schupp and Wagner, 2010).

Large national panel surveys are particularly suitable for collecting biomarkers because (1) biomarker data can be linked to multiple waves, (2) panels enable repeated measurements which can reduce measurement error (3) the sample sizes of many household panel studies create more opportunities for inference than small individual studies (Schnell, 2009).

The aim of this paper is (1) to study the feasibility of collecting cell material for genetic analyses with normal interviewers rather than medical personnel, (2) to evaluate the DNA quality of a new technique for collecting biomarkers, and (3) to analyze the selectivity of response. In section 2, we give a brief overview over related studies. In section 3, we explain the general methodology and introduce a new technique to obtain non-invasive cell material for genetic analyses. This new technique provides a DNA quality superior to that from the common cheek swap. In section 4, we report the results with a focus on response rates, selectivity, and the quality of the new biomarker technique. Section 5 concludes.

2 Overview of related studies

A comparison of response rates among studies is difficult for several reasons. Studies differ in target population, topic, and survey mode, all of which can have a substantial impact on response. Further, attrition in a cross sectional survey or

Contact information: Matthias Schonlau, RAND Corporation, Pittsburgh, USA, e-mail: matt@rand.org

the initial wave of a panel survey is generally higher than attrition in subsequent waves. Nonetheless, it is useful to consider the experiences from other studies.

Several surveys in multiple countries surveys have started to collect biomarkers. In the Danish LSAT survey of the oldest old, the collection of a blood sample had no affect on response rate (Christensen, Bathum, and Christiansen, 2008). The cooperation rate of the first wave without the collection of blood samples was 77%; the cooperation rate of the second wave with the blood sample was 81%.

In the U.S. the Health and Retirement Study (HRS) – a survey mostly conducted by phone – has an enhanced face-to-face interview that includes salivary DNA and blood spots. Among interviewees 94% consented to physical measures, 82% to the DNA sample and 81% to blood spots (Weir, 2008). In the ADAMS pilot study (Weir, 2008) only 11 out of 856 refused a cheek swap. However, in a diabetes study that used a self-administered blood test kit the participation rate was only 52%.

In the Taiwanese longitudinal Social Environment and Biomarkers of Aging Study (SEBAS), face to face interviews were conducted with the elderly (≥60 years) and pre-elderly (≥50 yrs, <60 yrs) followed by hospital appointments during which the respondents gave urine and blood samples (Chang, Glei, Goldman, and Weinstein, 2008; Weinstein and Willis, 2001). The response rate for the interview was between 91% and 93%. All but 10 of 1497 persons of those interviewed subsequently provided both the blood sample and the urine sample (Chang, et al., 2008).

A follow-up study to the National Collaborative Perinatal Project (NCPP) with participants in their mid 40s in New England, USA, found that blood samples could be obtained for only 70% (430 out of 618) of those interviewed (Gilman, et al., 2008).

Additional longitudinal studies that have collected biomarkers include the Health Survey for England (HSE), Midlife in the United States (MIDUS), National Social Life and Health (NSHAP) and the Wisconsin longitudinal study.[1] The "Chicago Core on Biomarkers" provides an over view over biomarkers in population settings.

The ability to collect biomarkers with normal interviewers (not health professionals) as part of a country-wide panel survey reduces costs and greatly facilitates logistics. However, this issue has been under-explored. A small number of studies report collection of biomarkers by mail without an interviewer (Avendano, Scherpenzeel, and Mackenbach, 2011). Also, very little is known about how the collection of biomarkers in general affects selectivity. Respondents with more conservative risk attitudes were underrepresented when collecting blood (Roe, Haab, Beversdorf, Gu, and Tilley, 2009). Older age was associated with increased participation rates for a self-administered mail-in saliva collection (Avendano, et al., 2011).

We explore both these issues: One, the feasibility of collecting biomarkers with normal interviewers using a new non-invasive technique of collecting biomarkers and two, response rates and selectivity, including selectivity for risk aversion.

3 Method

As part of the larger German Socio-Economic Panel Study (SOEP) (Wagner, Frick, and Schupp, 2007) its fieldwork agency "TNS Infratest Sozialforschung" conducted a pretest panel by running three in-person interviews (CAPI) of a probability sample of German noninstitutionalized adult residents in 2005, 2007, and 2008 (Siegel, Jaensch, and Huber, 2009).

The three pretest panel waves contained questions about personality and everyday decisions and took an average of 30 minutes to administer. The initial pretest sample consisting of 2135 people was based on random route selection in which an interviewer starts from a fixed location and follows a predetermined path to identify a household and a random person within that household. For the 2008 survey, the third wave of the pretest panel, respondents were informed by mail that the upcoming survey would be different and especially important. During the in-person interview respondents were asked to supply a cell sample for genetic analyses. The sample was sent to the Laboratory of Neurogenetics (Department of Psychology, University of Bonn) for deep freezing and later DNA extraction and genetic analyses.

3.1 Non-invasive collection of cell material for genetic analyses

Extracting DNA from blood samples yields higher DNA quality than extracting DNA from other cell material. However, the use of blood samples in large scale surveys is problematic:

a) The attrition rate is high due to a selective drop-out of subjects with a blood- and injection phobia. This is a severe shortcoming if phenotypes should be assessed that are associated with this specific phobia (e. g. anxiety or neuroticism). The attrition rate is also higher than that of buccal swaps.
b) Collecting venous blood requires medical professionals because it is an invasive technique.
c) Invasive techniques increase the risk for infections or the lesion of veins.
d) Blood samples require immediate freezing of samples.

In order to avoid these disadvantages of blood samples, often cheek/buccal swaps (one buccal swap per subject) are applied. However, in our experience the attrition rate of buccal swaps is also high due to small DNA yields. We invented a method that combines the buccal swap with another already established method for collecting cell material, the mouth wash technique, and developed a protocol suitable for use in the laboratory as well as in field studies. Briefly, the method is as follows: Both participant and interviewer give written consent both on an information sheet as well as on a declaration form before collecting the buccal cell sample. Under supervision of an interviewer, participants rub for a duration of three minutes with a sterile Q-tip over the inner side of both cheeks in order to loosen the cells of the oral mucosa.

[1] http://biomarkers.uchicago.edu/studiescollectingbiomarkers.htm

Afterwards, participants rinse their mouth with Listerine© or an alternative brand of antibacterial mouth water (antibacterial mouth water can be purchased in most drug stores and supermarkets). Preferably, the mouth water should contain alcohol because alcohol increases the stability of the cells. But, alcohol is not a mandatory ingredient. After flushing the mouth for one minute, participants spit the fluid containing buccal cells and Listerine into a meat-juicecollector (MJC, SARSTEDT, Germany). The MJC is a vessel with a removable funnel on top. After spitting the cell fluid into the MJC, mucosa cells clinging to the cotton bud of the Q-tip are discarded into the fluid by pressing and rubbing it against the inner surface of the MJC. Then the funnel is removed from the MJC and the tube is sealed by a plastic plug. The MJC tube is stored in a freezer for long term preservation. Storing the MJC at room temperature for several days does not degrade the quality of the cell material. Therefore, it is possible to ship the sample without dry-ice. In order to extract DNA from the cell material, the samples are first defrosted at room temperature for about 15 minutes. Then the tubes of the MJC are spun for 4 minutes at 4000 rpm. Half of the supernatant (about 2 ml) is discarded and the remaining fluid must be vortexed or, preferably, mixed by hand (while not affecting the concentration of DNA fluid, hand mixing restores a homogeneous cell fluid more easily). This process results in a higher concentration of the cell material. The concentrated cell material is then stored in two 1.5 ml reaction tubes (one for use and the other as a back up). Afterwards, any commercial DNA extraction kit can be used to extract DNA out of the mucosa cells. Our preference is to use a robot for automated DNA extraction. Automated purification of genomic DNA was conducted by means of the MagNA Pure© LC system using a commercial extraction kit (MagNA Pure LC DNA isolation kit; Roche Diagnostics, Mannheim, Germany). From 200 μl of our cell material we obtain 100 μl of eluted DNA.

3.2 Survey Measurement

Willingness to take risks was measured with a question about self-assessed risk (Dohmen, 2010). The scale of the general question ("Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?") ranged from 0 to 10 with 0 corresponding to "not at all willing to take risks" and 10 "very much willing to take risks".

3.3 Analysis

We computed (non)response rates for all three waves. We explored differential nonresponse, i.e. selectivity, by regressing response (yes-no) on willingness to take risks and demographic covariates using a logistic regression. To adjust for nonresponse we computed nonresponse weights as inverse response probabilities.

The R-indicator is a recent measure of representativity (Schouten, Cobben, and Bethlehem, 2009):

$$R(\rho) = 1 - 2s(\rho)$$

Table 1: DNA quantity and absorbency at different wave-lengths as quality indicators

	Mean	StdDev	Min	Max
μg/ml	22.633	11.696	4.500	61.300
260 nm/280 nm	1.182	0.130	1.010	1.590
260 nm/230 nm	0.565	0.147	0.310	0.930
230 nm	0.086	0.052	0.023	0.233
260 nm	0.045	0.023	0.009	0.123
280 nm	0.039	0.020	0.007	0.098

where ρ is the response propensity and s(ρ) its standard deviation. The R-indicator can take values from 0 to 1. Low values correspond to a lot of variation in response probabilities, high values correspond to roughly equal response probabilities. It is possible to have a low nonresponse rate but a high R-indicator if the non response is not selective.

4 Empirical Results

4.1 Feasibility

The vast majority of the 63 reporting interviewers stated that the preparation and implementation of the saliva sampling was very easy (36%) or easy (59%). Only 5% of interviewers expressed some difficulties. Most interviewers also reported sealing and mailing the saliva tubes was very easy (44.9%) or easy (54.0%) and again few (1.1%) reported difficulties. A small number of interviewers mentioned that the procedure of collecting saliva was unpleasant and embarrassing.

About one third of interviewers (32%) found it "hard" or "very hard" to persuade respondents to participate in the collection of the saliva sample. In comparison to other surveys at the same institute this was a very large proportion. No such difficulties were reported for the survey instrument. Nonetheless, all but one of the 63 reporting interviewers indicated they would be willing to participate in the same kind of interview with the collection of saliva again. Once persuaded to participate, none of the respondents stopped participating due to discomfort. It is unknown whether any of the participants had tender buccal mucosa due to mucosal disorders (e.g., lichen planus), auto-immune disorders (e.g., Sjögren's syndrome) or exposure to head and neck radiotherapy. All but one saliva sample sent for genetic analysis (250 out of 251) were successfully analyzed and yielded data. One sample could not be analyzed because the tube was damaged during mailing after being sealed incorrectly.

4.2 Quality of the DNA samples

DNA yield and purity were measured by means of a BioPhotometer plus (EPPENDORF, Germany) in combination with a LabelGuard© Microliter Cell (IMPLEN, Germany). Table 1 presents the DNA quantity (μg/ml) and the absorbency of the DNA at 230 nm, 260 nm, and 280 nm including the absorbency quotients 260 nm/280 nm and 230 nm/260 nm. The results show that DNA concentrations are

Table 2: Demographic composition, Satisfaction and Risk aversion characteristics of samples in 2005, 2007, and 2008

	2005 Wave Regular Survey	2007 Wave Regular survey	2008 Wave collected biomarker
Male Gender	46.3%	47.8%	43.6%
age <30	20.5%	15.3%	17.6%
30≤age <40	15.2%	15.5%	14.8%
40≤age <50	19.7%	21.8%	22.4%
60≤age <60	15.0%	16.5%	16.8%
50≤age <70	15.5%	17.4%	18.4%
70≤age <80	10.7%	10.4%	8.0%
80≤age	3.5%	3.1%	2.0%
Region: East Germany	20.6%	22.7%	27.2%
German Citizenship	95.8%	95.9%	96.0%
Education≤9 yrs	37.7%	36.3%	34.0%
Education 10-12 yrs	33.5%	31.0%	32.8%
Education ≥13 yrs	28.9%	32.7%	33.2%
n	1012	490	250

Education categories correspond to the German three tier secondary education system (Hauptschule, Realschule, Gymnasium)

Table 3: Logistic Regression coefficients of response indicator on demographic variables and risk aversion. Regressions involve all respondents who agreed to participate for a wave

	2007 Wave Regular Survey		2008 Wave collected biomarker	
	Coefficient	p-value	Coefficient	p-value
Self-Assessed Willingness to Take Risks	0.00	0.936	0.10	0.017*
Male Gender	0.12	0.385	-0.38	0.049*
age <30	-0.96	0.000***	0.12	0.705
30≤age <40	-0.41	0.098	-0.23	0.495
40≤age <50	-0.24	0.303	-0.04	0.904
50≤age <60	-0.09	0.702	-0.12	0.716
70≤age <80	-0.25	0.353	-0.42	0.272
80≤age	-0.45	0.263	-0.75	0.220
Region: East Germany	0.29	0.084	0.49	0.032*
German Citizenship	-0.24	0.482	-0.13	0.793
Education≤9 yrs	0.06	0.719	-0.03	0.913
Education ≥13 yrs	0.46	0.007**	0	0.994
constant	0.40	0.330	-0.08	0.891

* p<0.05, ** p<0.01, *** p<0.001

high and that the purity is also good. The spectrometer results are supported by the fact that nearly all samples could be genotyped by means of Real-Time-PCR.

4.3 Response rates and selectivity

The demographic composition of samples in 2005, 2007, and 2008 is shown in Table 2. The initial 2005 wave contains slightly more women than men (53.7% vs. 46.3%). Nonresponse was 52.6% in the 2005 wave, 51.6% in the 2007 wave and 49.0% in the 2008 biomarker wave.

The logistic regressions (Table 3) of response among respondents of the previous wave show different selective nonresponse in both the 2007 and 2008 waves. In the 2007 wave East Germans and highly educated people were more likely to participate and the under-30 age group was less likely to participate. In the biomarker 2008 wave men were less likely to participate. Willingness to take risks was associated with a greater propensity to participate. Note this coefficient refers to a one point increase is in the 10 point range. Both regressions have modest pseudo R squared values (0.025 for 2008 response and 0.023 for 2007 response).

The R-indicator is 0.82 for 2007 (given 2005 participation) and 0.85 (also given 2005 participation). This suggests overall mild variability in the response probability.

5 Discussion

Our study demonstrates that medically inexperienced interviewers had no technical difficulties collecting a biomarker sample. However, a large proportion of interviews found it challenging to motivate respondents to participate,

possibly because non-standard interviews are more challenging and more time consuming. Nonetheless, the biological quality and quantity of the DNA samples was high enough for meaningful analyses. The newly introduced method for collecting biomarkers for genetic analyses in survey panel studies shows remarkable advantages in comparison to invasive methods. In addition, the new non-invasive method yields higher quality and greater amounts of DNA than the commonly used non-invasive buccal swap

We found that the collection of biomarkers did not lead to an increased nonresponse rate. However, different selection mechanisms were observed in waves with and without biomarker, but selectivity was mild in both cases as indicated by the large R-indicators. In the regular wave increased nonresponse was observed primarily among young respondents, respondents with less than 13 years of education and West-Germans. In the biomarker wave nonresponse was observed primarily among risk averse respondents, West-Germans, and male respondents.

These factors affecting nonresponse can be viewed in the framework of leverage-salience theory (Groves, Singer, and Corning, 2000). Leverage-salience theory explains how respondents make the decision to participate in the survey. Under this theory, different respondents weigh individual survey attributes (e.g. survey topic, length, attitudes towards the sponsor) in favor and against participation. Individuals assign different weights to attributes, i.e. the leverage of attributes varies. Whether the attribute is salient depends on whether it is known to the respondent at the time the decision to participate is made. Under leverage-salience theory, risk perception, attitudes and personal level characteristics can be viewed as factors that affect the leverage of individual attributes.

For face-to-face surveys our study adds to the growing body of literature that shows no evidence of increased attrition. This result does not necessarily extend to other survey modes because biomarkers typically require some sort of face-to-face contact that respondents in other modes are not used to. When attempting to switch a sample used to phone interviews to face-to-face interviews 10% of the sample was lost even without the collection of biomarkers (Weir, 2008:68). Nonetheless, the HRS is now collecting biomarkers. The feasibility of collecting biomarkers as part of Internet surveys has also been explored in subsamples of the Dutch LISS panel (Avendano, et al., 2011). However, participation rates were low with about 15% for blood and saliva subsamples, and 27% for waist circumference subsample.

Our study has several limitations. Risk aversion may have been more pronounced if collecting the biomarker involved a more invasive procedure as would be required, for example, for collecting a blood sample. It is conceivable that this would reduce a drop in response rates that we did not observe. Second, the stable response rate for the biomarker wave may reflect two contradictory influences on response rate that cancel each other out: a) response rate is lower due to the collection of a biomarker sample, and b) response rates in later waves may be greater because potential nonresponders have already left in earlier waves. In addition, participants in the later waves of the SOEP pretest sample were already familiar with both the survey request and the interviewer which may also contribute to a higher response rate. However, the response rate in the second wave is similar to that of the first wave. Therefore there is no reason to assume that without collecting biomarker an increase in response rate should be observed in the third wave. Third, collecting biomarkers may affect response rate only in the following wave. A respondent asked to allow the collection of biomarkers might not refuse immediately in the presence of an interviewer. However, the respondent might be less inclined to participate in the following wave. We cannot rule this out because there is no fourth wave.

It is important to tread cautiously to avoid alienating respondents in established survey panel studies. However, if the scientific reason to collect biomarkers is compelling, the benefit of collecting them may well outweigh the potential risk to the survey panel.

References

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: Advances in applied Methods and New Research Strategies.* New York: Routledge.

Butz, W., & Torrey, B. (2006). Some frontiers in social science. *Science*, *312*(5782), 1898-1900.

Chang, M., Glei, D., Goldman, N., & Weinstein, M. (2008). The Taiwan Biomarker Project. In M. Weinstein, J. Vaupel, & K. Wachter (Eds.), *Biosocial surveys* (p. 60-77). Washington, D.C.: National Academies Press.

Christensen, K., Bathum, L., & Christiansen, L. (2008). Biological Indicators and Genetic Information in Danish Twin and Oldest-Old Surveys. In M. Weinstein, J. Vaupel, & K. Wachter (Eds.), *Biosocial surveys* (p. 15-41). Washington, D.C.: National Academies Press.

Dohmen, T., Falk, A., Huffman, D., Schupp, J., Sunde, U., & Wagner, G. (forthcoming). Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences. *Journal of the European Economic Association*.

Gilman, S., Martin, L., Abrams, D., Kawachi, I., Kubzansky, L., & al., E. L. et. (2008). Educational attainment and cigarette smoking: a causal association? *International Journal of Epidemiology*, *37*(3), 615-624.

Groves, R., Singer, E., & Corning, A. (2000). Leverage-saliency theory of survey participation: Description and an illustration. *Public Opinion Quarterly*, *64*(3), 299-308.

Roe, B., Haab, T., Beversdorf, D., Gu, H., & Tilley, M. (2009). Risk-attitude selection bias in subject pools for experiments involving neuroimaging and blood samples. *Journal of Economic Psychology*, *30*(2), 181-189.

Schnell, R. (2009). Biometrische Daten. In C. Koenig, M. Stahl, & E. Wiegand (Eds.), *Nicht-reaktive Erhebungsverfahren* (Vol. 1, p. 45-60). Bonn: GESIS.

Schouten, B., Cobben, F., & Bethlehem, J. (2009). Indicators for the representativeness of survey response. *Survey Methodology*, *35*(1), 101-114.

Schupp, J., & Wagner, G. (2010). *Zum "Warum" und "Wie" der Erhebung von (genetischen) 'Biomarkern' in sozialwissenschaftlichen Surveys, SOEP papers on Multidisciplinary Panel Data Research* (Vol. 260). Berlin, Germany: DIW.

Siegel, N., Jaensch, A., & Huber, S. (2009). *Die Messung genetischer Grundlagen von Alltagsentscheidungen: Methodischer Aufbau und Ergebnisse von zwei Machbarkeitsstudien*. Munich: tns infratest.

Wagner, G., Frick, J., & Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP) – Evolution, Scope and Enhancements. *Journal of Applied Social Science Studies (Schmoller's Jahrbuch)*, *127*(1), 139-170.

Weinstein, M., & Willis, R. (2001). Stretching social surveys to include bioindicators: possibilities for the Health and Retirement Study, experience from The Taiwan Study of the Elderly. In C. Finch & J. Vaupel (Eds.), *Cells and surveys: Should biological measures be included in social science research* (p. 250-275). Washington, D.C.: National Research Council.

Weir, D. (2008). Elastic powers: The integration of biomarkers into the Health and Retirement Study. In M. Weinstein, J. Vaupel, & K. Wachter (Eds.), *Biosocial surveys* (p. 78-95). Washington, D.C.: National Academies Press.