Collection of biosamples and health data from people in a post-COVID-19 world

Meena Kumari, Institute for Social and Economic Research, University of Essex.

1. Background

ESRC is setting up a new Early Life Cohort study that will begin collecting data in the near future. Measurement of health using biomarkers is being considered, because health can be a research topic in its own right and offer the potential to aid explanations of social and economic change and outcomes. This report, commissioned by ESRC, outlines innovative methods for the collection of biomarkers, and seeks to understand which can be collected in the post COVID-19 world, how to include objective health measures collected by study participants in the proposed new cohort study, and to suggest the kinds of measures that would be useful. In order to fulfil a brief of consultation with the wider academic community, the comments of 7 experts (listed on page 12), particularly in relation to participant led biomarker collection inform this resultant report.

1.a. Why should the new Early life Cohort include the collection of biosamples?

The 2017 Longitudinal Studies Strategic Review recommended that ESRC establish a new Early Life Cohort (ELC) as soon as practicable. The purpose of this report is to discuss the feasibility of collecting bio-samples and other health data. While COVID-19 has made face-to-face contact complicated, the development and recent availability of vaccines may make this easier. However, face-to-face contact may remain uncomfortable for some people. It may also represent poor value for money and is burdensome and there has been a move to develop more remote or participant led sample collection. There are emerging innovative methods for collecting biosamples (such as blood or hair) and associated data (such as body composition or blood pressure) from people which may be suitable for use in an ELC study. Some of these new methods appear to be efficient, effective, and exciting and non-burdensome for participants.

The ELC will build on a series of national birth cohorts, such as the National Study of Health and Development (NSHD, (1)), National Child Development Study (NCDS, (2)), 1970 Birth Cohort Study (BCS70, (3)) and Millennium Cohort Study (MCS, (4)) and more local birth cohorts such as Avon Longitudinal Study of Parents and Children (ALSPAC, (5)), Born in Bradford (BiB, (6)) and future studies such as Children Growing Up in Liverpool (C-GULL) and the Orebro Birth Cohort in Sweden enabling comparison of the lifecourse impacts of early life social, economic, biological and developmental factors in the pre- and post- COVID-19 era.

The ELC will predominantly be a social survey, which will be supplemented with measures of health and more specifically biomarker data for a number of reasons. For social science and social, economic and health policy purposes in the post COVID-19 world, the arguments for inclusion of health measures in the new study are based on the impact of health on the economy and vice versa. Health and biomarker data should be collected to accurately measure, model and forecast social and economic change and consequences. Our aim is to suggest that social and health researchers should be involved in this new study to ensure that:

- health and biomarkers are well measured for the purpose of understanding life course processes across the socio-economic spectrum

- health can be measured, and usefully modelled and forecast by epidemiologists and health economists using these data
- health and biomarker data in this study have independent epidemiological value

1.b. Sample constraints and opportunities

The new ELC will be the latest in a series of birth cohorts in the UK. We should consider the collection of biomarker data that would harmonise or enable researchers to capitalize on this richness of data to perform cohort and period comparisons. Further there are a series of geographically local birth cohorts that may also serve to act as complementary datasets. Thus, participants, schedules and measures in these studies have been reviewed and are used as points of reference here to enable comparison.

The cohort's design will be tested via a feasibility study with options for population over-sampling, potential questionnaire measures and biosamples will be evaluated. The collection of data in the post COVID-19 era will enable comparison of lifecourses before and after the pandemic and during the onset of potential future disruptions such as climate change.

1.c. What measures are we discussing when we talk about biosamples and other health measures?

We collect biosamples, such as blood and saliva, to assess biomarkers. The National Institute for Health has defined the term biomarker as 'a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention' (7). In the context of social surveys, the term biomarker is often used to encompass a variety of biological data from clinical risk factors for disease to factors of relevance to population health and well-being. Occasionally, the latter can be termed biomeasures, which include physiological measurements such as blood pressure but also wider function such as cognitive function (with 'biomarker' reserved for blood based markers). Here biomarkers will be used as an encompassing term that includes blood-based markers of biological processes and broader objective measures such as blood pressure and adiposity.

2. Choice of biomarkers

The study of the large and representative sample of the new birth cohort should include health measured in some detail, using the indicators already shown to be of predictive value. The choice of biomarkers should be hypothesis driven and should involve expert advice from appropriate early life health scientists. As discussed, the need to harmonize with measures with the earlier birth cohorts and current initiatives should also inform the biomarkers selected. However, decisions should also be cognizant of new data types and areas of research that the new birth cohort should capture. The health and biomarker data will, together with the detailed data on social and economic circumstances and change, provide not only the potential to address essential policy questions, but could also be of value to the health sciences.

Here we propose that biomarkers should be collected in the first and subsequent waves of the new birth cohort and we shall consider:

- who biomarkers should and could be collected from,
- when biomarkers should and could be collected,
- which biomarkers should be collected,
- which other measures should be collected to complement the recommended biomarkers.

2.a. Who could we collect biomarkers from in the new ELC?

With respect to participants, the birth cohorts have not restricted data collection to the infant cohort member and it is recommended that protocols in the new ELC are administered to cohort and family members. A number of studies in the UK and abroad (for example, ELFE (8) in France) have collected health data, including biomarkers, from family members. A list of the family members from whom samples are collected can be found in table 1.

Table 1: Health, biomarker and tissue collection from household members in British and other Birth Cohorts

	MCS	BCS70	NCDS	NSHD	ALSPAC	BiB	ELFE
Mode	F-to-F	Clinic	Clinic	Clinic	Clinic	Clinic	Clinic
Cohort members							
health	√	√	√	√	√	√	√
Biomarker collection	Anthropometry Actigraphy	Anthropometry Blood pressure Physical functioning	Blood pressure Anthropometry Physical functioning	Blood pressure Anthropometry Physical functioning	Blood pressure Anthropometry	Blood pressure Anthropometry (DEXA subset) actigraphy	
Tissue samples	Saliva Teeth	Blood	Blood Saliva	Bloods Saliva Spinal fluid* Imaging	Saliva Bloods Teeth Nails Urine Hair DEXA	Bloods Urine (subset)	
Mother							
health	√	√	√	√	√	√	√
Biomarker collection		Self reported: height, blood group and Hb data Blood pressure	Recalled prepregnancy weight Blood pressure	Measured height		DEXA, (subset	
Tissue samples	Saliva sample				Hair* Nail * Bloods	Blood	Serum Urine Hair
Father							
health	√				√	√	
Biomarker collection				Reported height		DEXA, (subset	
	Saliva sample				Hair* Bloods Urine	Blood	
Sibling	√					√	
health	√					√	
Biomarker collection		Older siblings reported birthweight		Younger sibling reported birthweight (+6y)		Which ones?	

Hair and nail collection at ages 3 and 4 in ALSPAC were collected by cohort member mother and partner. DEXA=Dual-energy X-ray absorptiometry to calculate adiposity. Most measures of parents in the birth cohorts collected in first wave of data collection

Cohort member: There are a number of measures that could be collected to complement and enable comparison with earlier studies, for example birthweight, weight and length of the infant are routinely collected in the earlier national birth cohorts and the local birth cohorts. Many of the studies that collected these data did so in face-to-face modes (NSHD, NCDS, BCS70) but others were collected by mother's report by questionnaire. In case of MCS, mothers were interviewed when their babies were aged 9 months and gave recalled information about the cohort member child.

Parents: Most birth cohort studies capitalize on the study design that necessitates involvement of the family members and collect data from parents, in particular the cohort member's mother. Thus, ALSPAC and MCS have collected biosamples from parents, in particular saliva sampling for the extraction of genetic material that enables genome studies in these cohorts.

Siblings: there has not been a tradition of collecting biomarker information from siblings to obtain a complete picture of the family in the birth cohorts. However, some data have been collected, for example, 1946 cohort asked about birthweight of babies from pregnancies subsequent to recruitment to the study. This appears to be an interesting omission, given the role of family dynamics on the development of children and future social and psychological health. The collection of biosamples could represent an interesting addition to what data are available in the other birth cohorts.

2.b. When do we collect biosamples?

To highlight a health component to the study at its outset may improve response rates and will help to encourage cooperation with subsequent biosample collection. Information collected at the first interview may also provide a sampling frame for sub-studies where numbers allow.

Table 2: Age at data collection waves in birth cohorts

	MCS	BCS70	NCDS	NSHD	ALSPAC	BiB	ELFE
Year 1	9 months	Birth	Birth	Birth	Birth, 6-8m 12m	Birth, 6m (subset) 12m (subset)	Birth 2m 3-10m 1y
Pre-school	3y			2y 4y	18m 31m 31m 43m	18m (subset) 24m (subset) 36m (subset)	2y 3.5y 4-5y
School age	5y 7y 11y	5y 10y	7y 11y	6y 7y 8y 9y 10y 11y	5y focus@7 focus@9 focus@11	Reception Primary school survey Actigraphy	5y

What sort of data collection schedule is being envisaged? It is yet to be decided but assuming a schedule similar to the Millennium Cohort Study, and not ALSPAC or BiB would mean that there would be no collection of biomarkers at birth but a data collection in the first year of life that would

enable a new study to be comparable to MCS. It is important to collect a biomarker in the first wave of data collection to sign post the study as a study of both social and biological factors. Thus a wider range of biomarkers can be collected in a second wave, when assuming a schedule similar to other birth cohorts, cohort members are aged 2-4y.

2.c. What can be collected?

Health and biomarker data collection should aim to harmonise with the earlier and current birth cohorts. With respect to health these cohort studies have collected a wide variety of health data throughout the lifecourse. These have ranged from recalled data such as birthweight reported in the MCS when the cohort member infant was 9months old, to actigraphy data that collect objective assessment of physical activity and, increasingly, sleep behaviours in children and adults. In table 3, we list which measures are routinely collected in the early waves of the birth cohorts and how these data have been collected.[1]

Table 3: early life biomarker data collected from cohort member child in birth cohorts

	MCS	BCS70	NCDS	NSHD	ALSPAC	BiB	ELFE
Birthweight	Recall by interviewer	Clinic	Clinic	Clinic	Clinic	Clinic	Clinic
Anthropometry	Recall prompted by PCHR Interviewer Home visit	Self report Home visit	Home visit	Home visit Clinic	Clinic	measured or PCHR	Measured and report by phone interview
Tissue samples	Interviewer home visit	Nurse home visit	Nurse home visit	Interviewer home visit Clinic data collection	Clinic data collection	Clinic data collection	Clinic data collection
Physical activity/sleep	Interviewer home visit	Nurse home visit	Nurse home visit		Centralised clinic	Centralised clinic	Centralised clinic

PCHR =personal child health record, sometimes called the 'red book'

3. How might we go about collecting biomarkers?

Sample collection needs to consider a number of challenges, in particular due to the longitudinal design of the new birth cohort. Such challenges include:

[1] For further details of measures of health see: https://www.closer.ac.uk/wp-content/uploads/CLOSER-Resource-Report-3-Bio-Measures-in-Longitudinal.pdf and for biomarkers see https://www.closer.ac.uk/wp-content/uploads/A-guide-to-the-biomarker-data-in-the-CLOSER-studies-FINAL.compressed.pdf

● Consistency of biosample collection and relevance of measures over time and at different ages

● The potential of feeding back results from biosample collection and panel conditioning from this feedback

● Choices of when and who to ask for sensitive and invasive data

● Cost

This paper is commissioned to understand how to include objective health measures collected by study participants, to examine whether these collection methods are tolerated and whether data collected are comparable to the measures collected in the earlier birth cohorts to enable comparative analyses.

3.a. What are the best methods?

As can be seen in table 3, many studies have conducted biosample collection in a centralised clinic or in the home (eg. NCDS, *Understanding Society*) but recently, this has not been possible necessitating a move to understand how and whether participants can collect their own biosamples. In an experiment in the *Understanding Society* Innovation Panel (IP12) that compared biosample collection by a nurse or interviewer in a face to face visit in the home with web based participant led collection, the following biomarker collections were tested in adults: blood pressure, hair sample collection (additionally collected from youth), dried blood spot collection, and interview administered to participants about stool sample collection. It was observed that participants are equally likely to take part in the survey if offered interview by nurse, interviewer or by web. Once included, successful biosample collection rates were higher for nurses and interviewers than web administered collection. Overall, usable samples were 60% in nurse administered collection compared to 30% in interview or web-administered collection. However, participant led sample collection was not biased by age, sex or socio-economic factors. Participant led biosample collection in adults is well tolerated as over 80% of participants that collected their own biosamples reported that they would participate in another sample collection if asked. The experiment has highlighted practical and cost issues that need to be considered. Findings suggest that the offer of feedback had a large and positive impact on response rates (Benzeval et al., in preparation). Timely feedback of results has important implications for a study as the generation of timely data is often expensive and there needs to be some consideration of the manner of feedback and the support participants might require in order to positively receive and utilise the feedback.

These results indicate a number of things:

a) participant-led biosample collection is well tolerated
b) a limited biosample collection in the first wave of a study may be a positive experience that potentially enables a more expanded biosample collection in subsequent waves
c) despite possible cost implications, feedback of results from biosamples should be offered, organised and supported to optimise response rates
d) as this was a study of adults, further development work is needed to examine parent led collection of biosamples from young children

3.b. What can be collected in the ELC?

The following biosampling strategy from the infant cohort member and other members of the household would provide a unique data source that would enable exciting and innovative research

within the study and also enable researchers to combine and compare with a variety of studies in the UK and beyond. Feasibility studies should be conducted to test the participant led collection of the following samples from the following participants.

3.b.1. Collection of health and biosamples in the early years: cohort infant

Anthropometry: Early life growth and growth trajectories are predictive of future health (9). Anthropometry was collected in all previous studies (table 3). The studies that have collected biological data in the early years are geographically specific and have done so in a clinic setting, which dependent on the chosen schedule of data collection may not be possible. It is recommended that the following measures are collected from the infant in the first wave of the study. A number of studies collect information on infant growth from routinely collected data, such as the infant personal child health record 'red book'. These measures include infant length and weight and provide information on growth trajectories in addition to clinical data such as immunisations. For example, the Millennium Cohort Study interviewers in the baseline sweep, when the infant was aged 9 months, asked parents to refer to the book to report the infant's last measured weight and immunisations.

When preparing for this wave, experiments had been performed to understand whether interviewers could collect 'red book' data but this was found to be time consuming and interviewers reported that respondents were more up to date and knowledgeable than the data in the red book (Prior). Given this, it suggests that there is potential for remote collection of growth trajectories to be more successful than this attempt at interviewer mediated collection. It is recommended that experiments should be conducted to examine the collection of height and weight data through either parental data entry or photographs of the appropriate pages.

Microbiome: It is interesting to characterize the 'microbiome' (microbiota inhabiting the gut) as increasingly it has become evident that there may be associations with illness, health and behaviour (9). This research is in its infancy and it is not clear what explains differences in the microbiome between people, how the microbiome develops over the lifecourse, particularly in healthy populations. The collection of a biosample to measure the 'microbiome' would be capitalising on recent advances in research and serve as an enhancement in comparison to earlier studies. There are now methods that should enable participant led collection of infant stool samples. It is likely that stool sample collection using established methods (eg DNA Genotek) would be feasible as attested to in preliminary work in Sweden (Montgomery, Orebro birth cohort) and from the stool sample collection in the ELFE study (Marie-Aline Charles, ELFE). *Understanding Society* conducted a focus group to investigate the acceptance and feasibility of methods to collect stool samples and found that participants were open to this collection, once they understood the scientific rationale of the collection (10). Methodological development is required to ensure a robust collection. A number of factors are considered to impact measurements in children and adults (eg breastfeeding and diet) and thus appropriate interview data would need to be collected at the same time as the sample.

Collection of a number of biosamples should be deferred to a second or later wave of data collection:

Hair sampling: New methods have been developed to measure biomarkers of stress. Of these development of methods to measure in hair are of interest as they are well tolerated and, compared to earlier methods, easy to administer (11). Experience in ELFE suggests that hair sampling from

cohort member infants, for the assessment of steroids such as the stress hormone cortisol, should be deferred to collection at aged 2-3y when children are more likely to have enough hair for hormone assessment as infants under the age of 1y do not consistently have sufficient hair for sampling (Marie-Aline, ELFE). In ELFE, this sample collection was successful when cohort members were aged 3.5y. The successful collection of hair samples in children in different settings has been described previously (12)

***Saliva sampling* for DNA analyses**: a number of social surveys have collected saliva samples for genetic analyses, for example a review of CLOSER studies can be found here CLOSER-resource-Harmonisation-of-strategies-for-exploitation-of-biological-sample.pdf. Thus collection of saliva samples to assess DNA would provide genetic data enabling trios analyses as in MCS, which collected saliva samples at age 14y using the Oragene DNA Kit. However, discussion with a number of studies that have collected biomarkers from early ages suggest that tissue sample collection may be difficult from infants younger than aged 1y. Thus, saliva sampling should be deferred to a subsequent wave. Similarly, it is likely that blood spot collection will not be tolerated by parents (Marie-Aline ELFE), Indeed it is suggested that blood sampling should not be attempted until the cohort members are much older as response rates are likely to be low and the participants characteristics biased by social factors.

3.b.2. From other members of the household:

Collection of samples and tissues from other members of the household would provide invaluable information of the family and the social and biological interactions in the early life environment of infant and it social and psychological development. Further, the family design available in the feasibility study would be an enhancement of earlier studies, for example genome wide studies which have encountered population stratification (confounding) issues when examining social outcomes.

Parents

Hair samples: Collection of hair samples from adults would enable the assessment of stress hormones in parents which in addition to questionnaire data assessing adult mental health would provide an insight into the psychological environment in the household. In *Understanding Society*, despite a lower response rate, compared to nurse collected hair, quality of hair sample collection by participants was high and not biased by age or educational attainment (Kumari et al., in prep). A potential problem for a potential birth cohort, is that it may be easier to collect in mothers than in fathers, where the response rate in the *Understanding Society* Innovation Panel was lowest in men due to a lack of hair.

Blood pressure: Analyses suggest that self collected blood pressure measures are lower than nurse measured but correlation is very high (Benzeval et al., in prep). However, the participant response is lower and biased by the health of the participant, as those with existing hypertension were more likely to return a measurement that those without (Benzeval et al., in prep).

Blood spots: Dried blood spots are feasible and until recently there were limitations to what can be measured in the spots, due to the limited lab capacity in the UK (Kumari et al., in prep). However, there are now new options with commercial companies now established that collect samples into blood tubes. Early evidence from *Understanding Society*, which collected blood spots before and

during the pandemic suggests that participants are more likely to agree to and return blood spot samples in the pandemic. Laboratories, organisation and capacity have been developed during the pandemic (for example, THRIVA), as required for remote collection. This sampling would enable the collection of blood samples to measure a variety of analytes, such as cholesterol, C-reactive protein, glycosylated haemoglobin and vitamin D and extract DNA, which would provide genetic data to conduct analyses in trios (two parents and one child) as has been conducted in MCS.

Stool sample collection: Data from the *Understanding Society* focus group indicated that participants would tolerate stool sample collection. The focus group was largely women and additional development work would be needed to understand whether there are gender differences in this type of collection (10).

Physical activity: Many studies have successfully collected actigraphy data to assess physical activity in adults (for example, BCS70 (14) and the Whitehall II study (15), with data suggesting lower participation in those with poor health. Methods are continually being developed and these types of data can also be used to calculate additional behaviours such as movement and sleep if supplemented with appropriate questionnaire data. However, these studies have been administered in face-to-face interviews and it is unknown how well these data can be collected using remote methods. Feasibility studies would need to be conducted to understand response rates to the request to administer this collection, the practicalities of data collection, and whether good quality measurements can be collected by participants themselves.

Height and Weight: There are currently no methods for the remote assessment of height and weight. These measures are collected by self report in a number of studies, including BCS70 and NCDS.

Siblings:

To be collected from older siblings in the household:

Hair sampling: Assessment of steroids would enable research into family dynamics and biomarkers of stress in the household (12). As discussed in the ELFE study, the collection of hair has found to be successful in children aged between 2-4y. However, in IP12, hair sample collection was well tolerated but response rate in those aged 10-14y was low (Kumari et al., in prep) suggesting that this collection method in much older siblings would need some development to improve response rate.

Saliva sampling: It is feasible to collect saliva samples from infants for subsequent genetic analyses as sampling kits and methods are available to enable the collection of saliva from young children.

Anthropometry**:** as with the cohort member, it is recommended that the collection of height and weight trajectories for younger children from the red book should be investigated.

Table 4: Suggested biosample collection schedule for the new ELC

	Baseline data collection	Subsequent data collection
	Cohort member under 1y	Cohort member between 2-3y
Cohort member	Recalled birth weight Anthropometry from red book Stool sample	Saliva sample Hair sample Stool sample

		Anthropometry from red book
Siblings	Saliva sample Hair sample Height* Weight*	Hair sample Height* Weight*
Mother	Blood pressure Blood sample Hair sample Stool sample Height Weight	Blood pressure Weight Hair sample Stool sample Actigraphy
Father	Blood pressure Blood sample Stool sample Hair sample Height Weight	Blood pressure Weight Hair sample Stool sample Actigraphy

*Height and weight would be collected from red book or self report as age appropriate.

4. Costs

Excluding sundries such as postage and packaging and development of information material, approximate costs for the following:

Microbiome – sample collection is £10 per sample and metagenomic measurements are approximately £100 per sample

Hair sample- sample collection is £2 per sample and steroid hormone measurements are approximately £30 per sample

Blood spots- using the example of 'THRIVA': Measurement of a panel of 5 analytes approximately £90 per sample

Extraction of DNA and measurement of methylation: £200 per sample

5. Ethical issues including any risk of harm to respondents.

There are a number of ethical issues to consider in the collection of biomarker data from participants. In a longitudinal study such as the focus of this paper, it is not possible to foresee all possible research uses at the outset. Thus, the study should seek to obtain consent for future use of biosamples. However, this is likely to be an iterative process with participants as research develops and participants are resurveyed. Guidance and practices have evolved over the life time of the cohorts and changed as a result of legislation such as the UK Human Tissue Act (2004) and the adoption of GDPR. Within GDPR, genetic data are considered sensitive and the sharing and use of these types of data by the research community have been dealt with through the use of data access committees. It is unclear how this will develop over time. This is particularly salient in the proposed sample collection, given the focus on families, where particular forms of or combinations of data may raise risks of disclosure.

6. Future proofing:

The suggested biosampling schedule should enable important and interesting research within the first wave of data collection. However, it is important to establish the survey as one interested in health and biomarker collection at the outset to help with future waves of data collection. However, it should be noted that decisions made at baseline have the potential to limit future analyses. For example, to conduct longitudinal epigenetic analyses, saliva sampling at baseline limits future sample collection to saliva as epigenetic data are tissue specific.

The study should also future proof in order to capitalize on developments in methods and research. In order to do this, samples in early waves should be collected and stored to optimize their future use. Thus, consideration should be given to the storage of stool samples following sample processing. Hair samples should be processed and condensates aliquoted and stored. Similarly blood samples processed, aliquoted and stored for future analyses. These processes necessitate an upfront and on-going cost for sample handling and storage that needs to be considered at the outset.

7. Additional data to complement biosamples:

We propose that additional health data should be collected directly from adult respondents in the household on:

- current self-reported state of health including information about illness and disability, and mental health screening questions
- current care of own and children's health, including current and children's past prescribed medication
- current health related habits of smoking, exercise, diet, alcohol consumption and sleep duration and quality
- care of others

Participant consent to NHS record linkage should be obtained to enable linked to mortality records, cancer registration, maternity, hospital and primary care data.

References

1.Wadsworth M. The origins and innovatory nature of the 1946 British national birth cohort study. Longitudinal and Life Course Studies 2010; 1; 121 - 136.

2. Power C; Elliot J. Cohort profile: 1958 British birth cohort (National Child Development Study). International Journal of Epidemiology 2006; 35; 34-41.

3. Elliot, J and Shepherd, P. Cohort Profile: 1970 British Birth Cohort. International Journal of Epidemiology 2006; 35: 836-843.

4. Connelly, R and Platt, L. Cohort Profile: UK Millennium Cohort Study (MCS). International Journal of Epidemiology, 2014; 43: 1719-1725.

5. Boyd, A, et al. Cohort Profile: The 'Children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. International Journal of Epidemiology 2013; 42: 1-17.

6. Wright, J, et al. Cohort Profile: The Born in Bradford in Bradford multi-ethnic family cohort study. International Journal of Epidemiology 2013; 42: 978-991.

7. Biomarkers definitions working group. NIH, 1998.

8. Charles, MA, et al. Cohort Profile: The French national cohort of children (ELFE): birth to 5 years. International Journal of Epidemiology, 2020; 49: 368–369j

9. Eriksson, JG, et al. Early adiposity rebound in childhood and risk of Type 2 diabetes in adult life, Diabetologia 2003; 46: 190-194.

10. Sarkar, A, et al. The Association between Early-Life Gut Microbiota and Long-Term Health and Diseases. J. Clin Med 2021; 10: 459.

11. Abbassian, L, et al. The acceptability of collecting samples from Understanding Society participants for microbiome analysis.: Understanding Society Working Paper Series, 28th June 2019, Vol. 05.

12. Abell, J, et al. Cortisol from hair samples in a large observational cohort: The Whitehall II study. Psychoneuroendocrinology 2016; 73; 148-156.

13. Bates, RA, et al. Hair sampling for cortisol analysis with mother-toddler dyads living in low-income homes. Infant behaviour and Development, 2020; 61: 101499.

14. Senaratne, N, et al. Device-measured physical activity and sedentary behaviour in relation to mental wellbeing: An analysis of the 1970 British cohort study. Prev Med, 2021:145; 106434.

15. van Kuppevelt, D, et al. Segmenting accelerometer data from daily life with unsupervised machine learning. Plos One 2019;14: e0208692.

The following Panel of Experts were approached and provided feedback:

Professor Michaela Benzeval: University of Essex, PI: *Understanding Society*

Professor Marie- Aline Charles: University in Paris, PI: ELFE

Professor Rebecca Hardy: UCL, PI CLOSER

Professor Michelle Kelly Irving: University of Toulouse, PI: EQUITY

Professor Scott Montgomery, Orebro University Sweden, PI: Orebro Birth Cohort

Gillian Prior, NatCen, Lead Longitudinal studies

Professor Leonard Schalkwyk: University of Essex, Lead Genomic and Computational Biology