Biomarkers in the proposed UK Longitudinal Household Study

Meena Kumari, Department of Epidemiology, University College London (UCL)
Mike Wadsworth, Director, National Survey for Health and Development, University College London (UCL)
Margaret Blake, National Centre for Survey Research
John Bynner, Director, Longview
Gert G. Wagner, Berlin University of Technology (TUB) and Director, German Socio-Economic Panel Study (SOEP) Berlin

24 September, 2006

CONTENTS

Executive summary

1. The ESRC is setting up a new Household Panel Study (the UK Longitudinal Household Study, UKLHS) that will begin collecting data in 2008. In addition to measure of social, economic condition and family dynamics, measurements of health using biomarkers is being considered. The purpose of this report is to provide recommendations for the incorporation of health and biomarker information into the UKLHS.

2. To meet the requirements of the brief a team was formed comprising social science and biological science expertise from Britain and overseas in the design and conduct of longitudinal surveys, including household panel studies. In addition, in order to consult more widely among the medical and health sciences community, 18 experts were sent the recommendations outlined here and their comments informed the final report.

3. The scientific rationale for the inclusion of biomarkers includes new social and economic policy questions as a result of changes in social mores, and changes in social and economic structures brought on by changes in population health, for example increases in life expectancy and increased prevalence of such conditions as asthma and obesity. At the same time the large sample size, wide age range and family-based features of the UKLS design offer unique research opportunities for medical science. .

4. The large age range allows the capture of developmental processes and decline. Biomarker assessment will thus vary in frequency depending on the stage reached. The highest frequencies will be needed when the rate of change is largest in childhood (growth) and in old age (decline).

5. We recommend the collection of a minimum set of health measures at the outset of the study. For the generation of biomedical resources for longer term scientific uses we recommend the formulation of hypothesis-driven studies as a basis for data specification, so that full scientific justification is available for the inclusion for any given biomarker. We recommend that the collection of biomarker information to support such studies be delayed to the second or third wave of data collection.

6. We recommend that cognitive function and a short battery of measures to assess health, disability and health behaviours (smoking, alcohol intake, physical activity, diet) is administered by the interviewers at the outset of the study. These will be age- specific in varying degrees. These measures will be collected intermittently with age dependent frequencies.

7. We recommend the biomarker information be collected by means of a nurse visit to the participant's home. These measures would be of: blood pressure, anthropometry, functioning and the collection of a blood sample. These measures

will be the same across the life span but will have differing 'age of onset' and will be collected intermittently with age dependent frequencies.

8. We recommend that a blood sample is collected for genetic studies as these make use of all the unique features of the study; the large size of the study, the large age range and the family design. The collection of genetic material in the UKLHS would make this a world class genetic resource.

9. Consent should be sought from participants for linkage to routinely collected medical data as available in medical records.

10. Organisational and leadership arrangements for the UKLHS should be available to ensure that capacity is available to undertake the specialised tasks that biomarker data collection entails

11. Arrangements should be made to ensure that full provisions for the preservation of confidentiality and data protection are in place

12. In conclusion, we recommend that the UKLHS should collect health and biomarker information as a resource for those examining the inter-relationships between health and social and economic outcomes. The size and family-based design of the study would also supply the basis for the creation of a unique genetic resource.

1. Background

The ESRC is setting up a new Household Panel Study (the UK Longitudinal Household Study, UKLHS) that will begin collecting data in 2008. The intention is to establish a nationally representative longitudinal study of a large number of households (40,000), in order to further understanding of social and economic change. Of particular concern will be education, work and retirement, income and wealth, and family dynamics. In addition, measurement of health using biomarkers is being considered, because it can be a research topic in its own right, but perhaps more importantly in the context of UKLHS, biomarkers offer the potential for better explanations of social and economic change. Biomarkers for both purposes are the subject of this paper.

To meet the requirements of the brief a team was formed bringing together social science and medical science expertise from Britain and overseas in the design and conduct of longitudinal surveys, including household panel studies[1]. The international longitudinal resources reviewed in the ESRC Strategic Review of Panel and Cohort studies were also drawn upon (longviewuk.com/pages/publications.shtml). In addition, in order to consult more widely among the medical and health sciences community, 18 experts (Appendix 1) were sent the recommendations outlined here and their comments have informed the resultant report.

Until recently, in social science panel studies, health has been measured by self-reported indicators, but there is now an increasing tendency to use more objective and specific measures, or biomarkers. The purpose of this paper is to consider the case for the inclusion of objective health measures in the proposed new British household study, and to suggest the kinds of measures that would be useful.

1.1 Foundations

The new study will build on the solid foundations established by the British Household Panel Survey (current n=9,000 households and 16,000 individuals) and some of the other leading household panel studies including the US Panel Study of Income Dynamics (PSID), the German Socio-Economic Panel Survey (SOEP), and the Australian Household and Labour Income Dynamics Survey (HILDA). Increasingly, in the interest of cross-study and cross national comparison, household panel studies of this kind are harmonising data collection topics and methods, including measures of health. There is, for example, the European Household Panel Study (ECHP), the Consortium of Household Panels for European Socio-Economic Research (CHER) involving 18 countries, and the more tightly targeted 'Cross-national Equivalent File' (CNEF) comprising the BHPS, PSID, SOEP,SLID and (soon) HILDA. Probably the most developed example of working to a common design and harmonized data is the Survey of Health, Retirement and Aging in Europe (SHARE) modelled on the US Health and Retirement Study and the English Longitudinal Study of Aging (ELSA), and involving 120 researchers in epidemiology, economics, psychology, and sociology from over ten European countries. All these collaborations recognise the increasing importance of

[1]Meena Kumari- ELSA, Whitehall II, NCDS age- 44 Biomedical follow-up; Michael Wadsworth, NSHD, NCDS age- 44 Biomedical follow-up; Margaret Blake, HSE; John Bynner, NCDS, BCS70, MCS; Gert Wagner, SOEP, Germany.

health data and the need to include biomarkers. In the case of such US studies as the 14, 322 strong longitudinal study of adolescent health over the ages 18-26 (Add Health) and the new 100,000 strong National Children's study (NCS), biomarker-related data collection goes further in taking measures of the physical environment, e.g lead content, atmospheric pollution, in which the children are born and grow up.

The need for harmonization highlights the importance of cross national research and the role of biomarkers (Burkhauser & Lillard 2005; Juster & Suzman 1995; Wolfson, 2006). Banks *et al* (2006) used data from HRS and the National Health and Nutrition Examination Survey (NHANES) and compared self rated health and biomarker data in ELSA and Health Survey for England (HSE). They found that while respondents in the United States report better self rated health than those in England, the objective assessments of health assessed by biomarkers suggest poorer health in the United States (Banks and Smith working paper). Cross-national comparative research also improves our understanding of how policies affect the choices people make because it allows examination of the effects of a much broader and richer mix of policies than is typically available in a single country. For example, Kenkel, Lillard, and Mathios (2004) document the substantial variation across the United States, United Kingdom, Germany, and the Russian Federation in the mix of taxes, workplace smoking bans, advertising restrictions, and warning label policies aimed at reducing consumption of tobacco. They note that the variation across countries in the different mixes of these policies can be used to study how decisions to smoke vary with policy mixes outside the range observed in a single country.

Thus, cross-national comparative research adds value to our understanding of basic human behaviour and to the analysis of how public policies affect choice. Funding bodies such as the National Institutes of Aging have been increasingly prepared to invest in the production of these large-scale cross national comparative datasets and to support the research based on them. For the UKLHS the implication is that design decisions, including the use of biomarkers, need to be taken with full awareness of comparable developments in other studies across the world and the potential for harmonization.

1.2 Scientific rationale for the inclusion of biomarkers

What are the arguments for household panels to collect health data, and in particular to use biomarkers? Increasingly such surveys concerned with social policy questions collect data on health, because changes in population health (e.g. both in terms of the large changes such as increased life expectancy, and the smaller but incrementally powerful changes associated for instance with antidepressant medication, and management of raised blood pressure) bring new social and economic policy questions. Equally, changes in social mores (e.g. increasingly late age at first birth, high rates of partner status change, changing social cohesion) bring new health policy questions (Halpern 2005), and changes in social and economic structures (e.g. unemployment) bring changes in health and well-being (Wilkinson 2006). Health is also of concern because epidemiological studies have shown that some health measures, together with indicators of the social and economic context and of the individuals' social relations, affect not only the likelihood of poor health at later ages (Barker 1998; Kuh & Ben-Shlomo 2004; Marmot & Wilkinson 2006)

but health and other biological facts can influence and predict social and economic outcomes (cf. for a brief overview Lillard and Wagner 2006; Biddle and Hamermesh 1994, Caspi et al. 2002, Ding et al. 2006, Knudsen et al. 2006, Moffitt et al. 2006, Smith 2004)

Health measures and biomarkers in general have value for a number of reasons. The life course perspective adopted in much social and medical longitudinal research emphasises the interactions across the different domains of life - education, employment, family, community, health and well-being - through which development across the different stages of life occurs. This underlines the point, considered in more detail later, that to capture developmental processes biomarker assessment will vary in frequency depending on the stage reached. The highest frequencies will be needed when the rate of change is largest in childhood (growth) and in old age (decline).

Thus health status becomes increasingly recognised as central to understanding life course processes across the socio-economic and age spectrum. For example, the health of the population aged 40 to 60 years is of concern in policy terms because of the cost of its current upkeep (e.g. in terms of managing ill health, and of control of risk by keeping blood pressure and cholesterol low, and maintaining functioning during the menopausal transition), and as an indicator of where health investment might best be made now. Current health also has value as a measure of the future health potential of that cohort in their later years (e.g. as measured in current obesity, physical and mental functioning, nutrition, and smoking and exercise habits). Current health of other age cohorts has comparable value. For instance, the future health potential of the cohort now in childhood can be measured by such indicators as their growth and development in early life, experience of illness, nutrition, and emerging habits in terms of smoking, exercise, alcohol consumption and drug use.

Although our prime consideration in this paper is to address social science needs for biomarker data and the requirements for interdisciplinary enquiry, the UKLHS design will also offer unique research opportunities to health scientists especially in genetics and epidemiology. The density of social information contained in this survey would be greater than most studies containing biological information. The size of the study and its household design would make a collection of biomarkers, in particular genetic material, unique and would make this a world class genetic study. The sample size is three or four times that of the existing longitudinal studies of health, which offers a unique opportunity to study genetic effects, and in particular gene/environment interactions, in relation to common outcomes, that have not usually been possible so far because of small sample sizes. The fact that data will be available on this scale about families is also of unique value for genetic studies. Family based studies are the accepted definitive method of controlling for population stratification (confounding) in genetic association studies (cf. Baker 2004) and reviewers and journal editors frequently require population studies be confirmed in family studies. Geneticists will also be attracted to the opportunity the study provides to investigate family traits in relation to illness and health function (e.g. blood pressure), health risk (e.g. raised blood pressure) and health protective outcomes.

Epidemiologists concerned with the interaction of the social environment with health will see unique opportunities for studies of within household/family effects (e.g. the health effect of unemployment in a key household member, of effort/reward assets from the family environment, of social mobility) in which the outcomes are likely to be specific illnesses, health risks (e.g. obesity, inflammatory status), health related habits (e.g. diet and dietary habits), and change over time in function (e.g. decline in respiratory function). Those interested in social outcomes can examine the interaction of sub clinical health processes and outcomes such as social engagement, participation in the workforce and retirement.

Apart from scientific merit, serving these different kinds of need has the attraction of interesting medical as well as social science research funders. However there is an important distinction between approaches to research resource production that needs to be borne in mind. Medical funders generally expect to see scientific programmes mapped out in advance, specifying the hypotheses that the data once collected will enable the proposers to test. Research resources for secondary use by others then become a spin-off from the programme. Support for new data collection in the social sciences tends to rely on consultation with the potential research community to ensure that measures of variables relevant to the most likely range of research uses are included in the data set. In this sense the exploitation of the research resource thus produced is all secondary analysis.

Our recommendation for biomarkers in UKLHS is to move more towards the medical science position, that is to say, the formulation of hypothesis-driven scientific studies as a basis for data specification, so that full scientific justification is available for the inclusion of any given biomarker. This approach will help both to minimise the expense of collecting this type information and to maximise the chances of funding from medical sources. These kinds of studies have a long lead time before data collection because of the organisation of data collection staff, laboratories and other data processing arrangements, and ethical review, as well as the grant application process. In order to ensure that the health measures will be of maximum scientific value and to give them the optimum chance of funding, it will be necessary for them to be planned as part of long-term research programmes designed to capitalise on the UKLHS assets described above. Such arrangements could not be ready for the first data collection.

We therefore suggest a minimum health data set to be decided on by the social scientists who will use the data, and to be collected as part of the first data collection. The minimum health data set will provide valuable information for social scientists concerned with the costs of health (in terms of NHS costs and spending in other sectors), and the extent and cost of preventive care (NHS prescribed and otherwise). The dataset will provide scales of state of current health and disability (self-rated). Their inclusion will also make clear to participants at the inception of the study that there is a health component to the study. To highlight a health component at the outset of a longitudinal study such as ELSA has been shown to improve response rates. In the UKHLS it will also help to encourage cooperation with the subsequent nurse or clinical visit. Maybe even the name of the study should include “health” (at least in the “field name” which will be visible to the respondents” (Socio-economic Life and Health in UK, for example).

Information collected at the first interview may also provide a sampling frame for sub-studies where numbers allow.

1.3 Sample constraints and opportunities

Details of the UKLHS design have yet to be decided. Most notably, it is unclear at this stage whether there will be 'over-sampling' to boost numbers of households in certain categories such as ethnic minorities and the populations of Scotland, Wales and Northern Ireland. The balance of argument in the other reports appears to us to be against boosting. Nevertheless, as a longitudinal household panel including all ages and quadrupled in size, the UKLHS will, uniquely offer the opportunity to compile data from cohorts of any age distribution, and to accumulate data about events such as divorce or hospitalisation over much shorter time spans than is possible in existing longitudinal studies (for examples of potential numbers see *ESRC Strategic Review of Panel and Cohort Studies* www.longviewuk.com/pages/publications.shtml).

The new longitudinal household panel study will have excellent statistical power for analysis of the subpopulations of major scientific and policy interest, and offers opportunities for cross cohort comparison and fine grained analyses of short and medium term developmental processes that complement and expand those available in the birth cohort studies. The UKLHS will also usefully complement the UK Biobank, ELSA and other resources (eg MRC National Survey of Health and Development, NSHD); National Child Development Study, NCDS; Avon Longitudinal Study of Parents and Children, ALSPAC; Millennium Cohort Study, MCS) because of its inclusive cross-sectional age structure (Biobank and ELSA both begin in middle life), family data (Biobank and ELSA are both studies of a sample of individuals), large sample size, and over time detailed life course data. The new study could therefore, in due course, enhance substantially the longitudinal health data resources currently available in the UK.

1.4 Choice of biomarkers

We argue that the study of the large and representative sample of the new UKLHS should include health measured in some detail, using the indicators already shown to be of predictive value. To recap, the choice of biomarkers should be hypothesis driven and should involve expert advice from appropriate health scientists. The need to harmonize with measures used in other countries should also inform the biomarkers selected. The health 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 would also be of value to medical science.

2 Research Questions

2. 1 Minimum health data to be collected by interviewer

We propose that data should be collected directly from respondents on:

- current self-reported state of health including information about serious illness and disability, and mental health screening questions such as the CES-D which can be administered from ages 14 yrs and over

- current care of own health, including current prescribed medication and non-prescribed forms of care, including diet supplements and other preventive health care
- current health related habits of smoking, exercise, diet and alcohol consumption
- care of others

The sample should also be identified as soon as practicable on the NHS Central Register (NHSCR), so that information on deaths and cancer registration could be received. Consent should be obtained from respondents for access to hospital and general practitioners' records.

2.2 Optimal set of biomarkers:

A number of issues such as respondent burden and financial cost of a data collection exercise such as the one outlined need to be balanced against the scientific information gathered. Our recommended measures can be classified by their cost but this may come at the expense of scientific questions that can be addressed. The least expensive panel of biomarkers is one which can be collected by the interviewer without the requirement for an additional nurse visit. However, these measures are limited and would not make optimal use of the longitudinal household design. We also believe these measures are not the ones that would help maintain response rate (see section 3.2 below). Interviewers have previously collected height, weight and cognitive functioning data (for example in SOEP which is an "all purpose" panel very much like UKLHS will be). It is envisaged that DNA from a cheek swab could also be collected by an interviewer but this would need to be tested as outlined in section 3.2 below.

We recommend that cognitive function and a short battery of measures to assess health, disability and health behaviours (smoking, alcohol intake, physical activity, diet) is administered by the interviewers at the outset of the study. These will of course be age-specific in varying degrees.

Cognitive function in particular raises major challenges for measurement continuity across the whole life span because of changes with respect to the:

- nature of cognitive function
- methods of measurement
- frequency of data collection needed
- interview time needed
- interviewee response to the assessment

All the birth cohorts (NSHD,NCDS, BCS70, ALSPAC, MCS) have used measures of cognitive function across the age span covered. For example, the MCS used the Bailey cognitive development scales and the British Ability Scales for preschool children (BAS II, NFER-Nelson). The age-34 BCS70 follow-up, included cognitive assessment of half the cohort's children using the age 3-5 and age 6-16 BAS II scales. NSHD has used the National Adult Reading test (NART.- NFER Nelson) to measure cognitive function through adulthood. The Whitehall II study assesses memory, AH4, phonemic and

categorical fluency (measures of executive function). ELSA includes the same measures of executive function and NSHD's measure of attention, amongst others.

The cohort studies also offer well-developed questions and standard scales on health-related and problem behaviours (e.g. the Cage scale of drinking behaviour), which would be starting point for measures to include in UKLHS. For school age children inclusion of the widely used age-graded 'Strengths and Difficulties Questionnaire' (SDQ) (Goodman, 1990) and the 'Rutter Behaviour Scales' Rutter, Tizzard and Whitemore (1970) also be desirable. Mental health measures through the late teens and adulthood include the General Health Questionnaire, GHQ, CES-D (which can be used aged 14+ yrs), the Malaise scale. Personality assessment comes outside our brief, e.g. the Maudsley Personality Inventory, MPI, (c.f. SOEP) and other measures of psychological attributes such as motivation, self-esteem and self-efficacy, but needs flagging up as a design issue to be resolved as recognised in Study 1).

We also recommend that consents be requested for linkage to the routine health data outlined in section 3.5 below. This should be followed in wave three by measures collected during a nurse visit. The biomarkers collected during the nurse visit have been measured in a number of international studies would therefore provide the optimal range of measures which could be used for cross cohort comparisons.

These measures would be:

- Cardiovascular measures such as blood pressure assessment
- Anthropometry such as height, weight and waist circumference
- Functioning assessment such as lung function and hand grip strength assessment
- Blood sample (serum and plasma) collection for storage for subsequent analysis, for example genetic markers.

We recommend that this battery of measures is staggered across waves as outlined in sections 3.3 and 3.4 below. This battery of measures is shorter and thus less burdensome to participants than in a number of recent data collections from large scale social surveys such as ELSA and NCDS but would provide an array of measures that make best use of the large age range, the household design and the longitudinal aspects of the UKLHS**.**

3. Methods

3.1. Collecting the biomarker data

Participants may be invited to a centralized data collection centre or visited in their homes. Inviting participants to a centralized clinic may reduce response rates in groups that are difficult to access in large scale epidemiological studies such as those with poor health or low socio-economic status. Visiting participants in their homes results in better

response rates but is more expensive and time consuming. In 2003 the Whitehall II study involved a comparison study of clinic and home visits by offering some respondents both a home and clinic visit and others two home visits to compare the results. This is because there is a possibility that an individual respondent's results may differ according to the setting in which they were taken. This would particularly affect blood pressure and pulse rate. The data from this comparison study have not yet been analysed, but such an analysis could usefully inform the choice of venue for any nurse component of the UKHLS. If results differ significantly between the home and clinic visits, this would argue against a mixed method.

3.2 Recommendations for core 'biomarker' information to be collected in the nurse visit

3.2.1 CARDIOVASCULAR RISK FACTORS

Diseases of the heart and circulatory system (cardiovascular disease or CVD) are the main cause of death in the UK and overall coronary heart disease (CHD) is estimated to cost the UK economy over £7.9 billion a year. Amongst other factors, risk of CHD is directly related to both systolic and diastolic blood pressure levels. The World Health Report 2002 estimates that around 11% of all disease burden in developed countries is caused by raised blood pressure. Raised blood pressure is determined by factors in both early life and in adulthood and by genetic factors and as such is we recommend that it is collected in the UKLHS.

a) Blood pressure and pulse assessment.

This uses a relatively simple piece of equipment (eg: Omron HEM 907 blood pressure monitor). The unit cost is about £375. This measure achieves high co-operation rates (99% on NCDS Wave 7). Information about adult blood pressure can be fed back to respondent during the nurse visit. Blood pressure has regularly been taken for children aged 4 and over in previous surveys although the results cannot be fed back and interpreted by the survey nurse for those under 16. Blood pressure assessment and feed back results serves to maintain response rate as it is a measure that the public 'knows'.

3.2.2 FUNCTIONING

Functioning tests assess the levels of physical ability and disability in the population. Disability has important social and economic implications at all life stages. In middle age the ability to work is compromised and development of disability has obvious implications for social other health functioning. In economic terms, disability is an important program in many countries, and one that until recently was growing rapidly over time. The number of people on disability programs is substantial, particularly among men and women in the age groups 45-64

Disability can be reported by self assessment or by objective measures such as lung function tests and physical performance measures such as hand grip strength, the former are commonly used in clinical practice to assess impairment due to chronic lung disease and asthma. Both these objective measures are taken in international studies such as ELSA and SHARE. Evidence from the 1946 cohort suggests that the development of these functions have substantial lifecourse influences (Kuh et al., 2006).

a) ***Lung function***

Studies measure forced expiratory volume, forced vital capacity and peak expiratory flow. This has been taken from children aged 7 and over on HSE and from older respondents on ELSA. Co-operation rates are generally high (98% of those seeing a nurse on NCDS Wave 7)

b) Grip strength

The grip strength measurement is an indication of upper body strength, and gives an objective, comparable measure of strength or frailty. It is measured with a gripometer which consists of a gripping handle with a strain-gauge and an analogue reading scale. Hand grip strength may represent a measure which can discriminate in younger age groups but this would need to be piloted.

3.2.3 ANTHROPOMETRY

Overweight and obesity increase the risk of CHD. As well as being an independent risk factor, obesity is also a major risk factor for high blood pressure, raised blood cholesterol, and diabetes. The adverse effect of excess weight is more pronounced when fat is concentrated mainly in the abdomen. This is known as central or abdominal obesity and can be identified by a high waist to hip ratio. The World Health Organization's World Health Report 2002 estimates that over 7% of all disease burden in developed countries is caused by raised BMI, and that around a third of CHD and ischaemic stroke and almost 60% of hypertensive disease in developed countries is due to levels of body mass index (BMI) in excess of the theoretical minimum (21 kg/m^2). Overweight and obesity are increasing rapidly. In England, the percentage of adults who are obese has increased by over 50% in the last decade. There has also been a steady increase in the prevalence of obesity in children.

a) Height and weight

These measures are easy to take and can be carried out by a trained interviewer. This means they can be collected during an interview and in waves when a nurse visit is not included. On a longitudinal survey adults would only need to be measured once but children could be regularly measured until they reach adulthood. Although height and weight are collected on many other surveys, this survey would be unique in measuring the whole household longitudinally, providing information about patterns within households over time.

b) Infant length

For children under the age of two nurses can measure their horizontal length, rather than interviewers measuring their upright height. In waves of the survey when there is no nurse visit, this information can be gathered from a book that mothers keep which records health visitor and other medical information about their children. The disadvantage of using this so-called 'red book' is that data are not collected in a standardised manner but the advantage is that information on height can be collected from under twos in every wave of the survey.

c) Demi-span

This measure is used for older people (usually 65 and over). Among older people measuring height can be quite difficult if the respondent cannot stand straight or is unsteady on their feet. Additionally, height decreases with age. This decrease varies from person to person and may be considerable. It is becoming increasingly important to have information about the health of older adults. Therefore an alternative measure of skeletal size, the demi-span, was developed which can be measured easily and does not cause unnecessary discomfort or distress to older adults. The demi-span measurement is taken in older adults and is the distance between the sternal notch and the finger roots with arm out-stretched laterally. Two readings are usually taken.

d) Waist and hip

This is a relatively cheap and easy measure to collect, requiring only a tape measure. It is regularly taken from children aged 11 and over. Co-operation rates are generally high (99% on NCDS and 97% on HSE among those seeing a nurse).

e) Body fat

The distinction between normal weight, overweight and obesity is conventionally made using BMI calculated from height and weight. However there is concern that some people, particularly men with high levels of muscle, may be classified as overweight using BMI even if they have low levels of fat. The use of bioimpedance to calculate fat levels using hand held body fat measurers or body fat scales could be considered. Up to now the body fat scales have not offered sufficient accuracy to be used as a replacement for conventional scales on the Health Survey of England. However, during the life of a longitudinal panel it is likely that technology may develop. Since obesity is a major and growing public health issue which clusters in households and is developing increasingly early in life it is important that the UKHLS adopts measures like this as soon as reliable technology becomes available.

3.2.4 BLOOD SAMPLING

Blood samples have been taken on a wide range of both and cross sectional and longitudinal surveys including NSHD, NCDS, ELSA, ALSPAC and Whitehall II. The equipment required is not costly. However this is the most invasive biomarker taken during surveys and achieves the lowest response. Comparing the NCDS with the Health Survey of England (HSE) suggests that co-operation for blood taking is much higher on a longitudinal survey than a cross-sectional survey (94% of those who had a nurse visit on NCDS compared with 75% of those who saw the nurse in HSE). Blood samples can be taken from children with parental consent (aged 11+ yrs on HSE and age 8+ yrs on low income dietary and nutrition survey (LIDNS)); however response rates are low in those aged under 10 and we recommend that blood sampling be limited to those aged 11 yrs and over with parental consent. For children giving blood samples it is standard practice to offer an anaesthetic gel such as ametop. We would recommend that blood samples be collected from children aged 11+ with parental consent.

A wide range of analytes can be tested, according to the research needs. Blood can also be stored for future analysis. Once a nurse is taking blood, 3 or 4 tubes can easily be taken without increasing the respondent burden or cost substantially. We would recommend storage of samples as we anticipate that high throughput techniques that allow for assay of an array of markers using small sample volumes will be developed during the life of the panel.

For most analytes blood samples can be sent by post to a central laboratory for analysis. However some analytes require blood which has been processed within an hour or two of it being taken. Examples of these analytes are Vitamin C & fat soluble vitamins & caretonoids and Water soluble vitamins (folate, homocysteine). These are of interest where diet and nutrition are the focus of study. UKHLS offers a good opportunity to research the longer term effects of diet on a household and so these analytes may be required in the future. However we are not recommending them at the outset given the increased resources needed to collect these data.

a) ***Use of blood samples for DNA***

Genetic material is now being collected from a large number of longitudinal studies that collect social and economic information (e.g. NSHD, ELSA, NCDS, ALSPAC, Whitehall II). Gene-environment interactions are of increasing interest in the explanation of behaviour and in the aetiology of disease (e.g. Moffitt, Caspi and Rutter, 2006). Genetic information can be extracted from blood samples. This is a convenient method to use on surveys where blood samples are already being given. On the NCDS and ELSA permission was sought from respondents who had given blood samples to extract the genetic information. On ELSA assurance was given that the genetic data would not be made available to insurers, mortgage applications, police records or for HIV or AIDS testing. On NCDS Wave 7 with 44 year olds 96% of those who gave a blood sample agreed that their DNA could be extracted and stored.

In order to obtain higher rates of co-operation for giving DNA samples it may be more effective to collect them using a mouth swab to collect sputum or cheek swab to collect cells (as is done for young people in the Add Health survey in the US). Although this may be less invasive and might give a higher response than a blood sample this should be piloted. It is possible that respondents will agree more readily for their DNA to be stored if it involves a sample which has already been taken from them. There is a possibility that a sample of sputum could be collected by an interviewer. However collecting sputum for a DNA sample via an interview, without other biomarkers being taken could potentially have adverse effects on response.

3.3. Timing of minimum health and biomarker data collections

As noted earlier, many biological functions show rapid growth in the early stages of life followed by stability and then a decline with age in later stages of life, including lung function, muscle strength and growth. Social and family factors in early life, together with genetic effects, influence the development of these measures and determine the 'peak' from which decline in

adult life is apparent. Comparable adult factors determine the timing and rate of adult decline. Questionnaire information would be collected at each wave or rotated throughout the data collection waves, but biomarker information collection should vary dependent on age group.

We propose that the first wave should include collection of non-invasive, questionnaire data on health (as described in section 2.1. above), and some aspects of that should be repeated in subsequent waves, as shown in table 1.

Table 1. Timing of minimum health data collections

	All ages **Self and parent reports of health and disability and mental health questions**	***Age 10 yrs and Above*** **Health habit Reports**	***Ages 16 and over*** **Medication, health care and Care of others**
Frequency of data collection	Annually	Waves 1,4,6,8,10	Waves 2,4,6,8,10

3.4 Timing of biomarker data collections

For reasons outlined in section above, biomarker data collection could be staggered across study waves to minimise respondent burden, field work practicality and cost issues and according to the hypotheses. Table 2 illustrates how that kind of timing could look. This schema ensures that no respondent participates in the whole health module (questionnaire and nurse visit) at any one stage to minimise burden.

Table 2 Timing of biomarker collection by age groups

	Measures of body dimensions	**Cardiovascular measures**	**Musculoskeletal & respiratory Measures**	**Measures of cognitive function**
Wave 1	Self report or red book			Baseline for all participants
Wave 2				

Wave 3	Baseline for half the cohort	Baseline for half the cohort	Baseline for half the cohort	16 and under
Wave 4	Baseline for half the cohort	Baseline for half the cohort	Baseline for half the cohort	
Wave 5				45 yrs+ 16 yrs and under
Wave 6	Under 16 yrs		10-16 yrs	
Wave 7				16yrs and under
Wave 8	45+ yrs Under 16	45 + yrs Under 16	45 yrs+ 10-16	
Wave 9				45 yrs+ 16 yrs and under
Wave 10	Under 16		10- 16 yrs	
Wave 11				All participants 44 and under
Wave 12	Follow up for half the cohort	Follow up for half the cohort	Follow up for half the cohort	
Wave 13	Follow up for half the cohort	Follow up for half the cohort	Follow up for half the cohort	45 yrs+ 16 yrs and under

3.5 Effects on response rates of including biomarker collection

Although biomarker collection is only just beginning in household panel studies, it has a long history in the birth cohort studies that also collect extensive social and economic data. Most biomarker collection in those studies is undertaken by specially trained nurse interviewers.

Although biomarker collection is only just beginning in household panel studies, it has a long history in the birth cohort studies that also collect extensive social and economic data. Most biomarker collection in these studies is undertaken by specially trained nurse-interviewers, without apparent adverse effects on response**.**

Agreement to nurse measures is higher in longitudinal surveys of individuals, particularly in later waves, than in cross-sectional surveys. Response rates for nurse visits were 88% (of those responding at Wave 2) in ELSA, 83% (of those responding at Wave 6 at age 43 years) in NCDS, and 83% (of those alive, not refused, and resident in England, Wales or Scotland at Wave 21 at age 53 years in NSHD. Notably, response rates in the age 46 telephone follow-up in NCDS appeared unaffected by the age 44 biomedical follow-up, and did not differ between those who had done and those who had not done the biomedical assessments.

3.6 Field work

3.6.1 Concurrent interviewing

HSE uses concurrent interviewing in which up to 10 adults and 2 children per household are eligible. Up to four members of the household are interviewed together, asking each short group of questions to each respondent in turn. This method could be used on nurse

visits in UKHLS. Although each respondent spends longer in their interview, the overall interviewing time is shorter than if each household member is interviewed individually. Furthermore, interviewing people together can aid recall. Based on HSE data we calculate that an adult nurse visit would take 30 mins and a child aged 5 and over, 20 mins. For a family of four, this results in a maximum visit that would last approximately 1.5 hours. However we are recommending such a visit only once approximately every 10 years.

3.6.2 Consent for linkage with national administrative data

In many surveys respondents are asked to consent to having their survey responses linked to information on central registers or records.

- NHS Central Register (deaths) (e.g. HSE, ELSA, NCDS, NSHD)
- NHS Cancer Register (e.g. HSE, NCDS, NSHD,ELSA)
- Hospital Episodes Statistics (e.g. HSE, ELSA)
- Birth and pregnancy records (e.g. MCS)

Agreement is generally high. These would be an important component of any health element on the UKHLS and consent should be sought at the first interview.

4. Research capacity and academic leadership

Effective implementation of a biomarker module in the UKLHS, and household panel studies more generally, would make new demands on the teams conducting them. This is because the research programmes embracing biomarkers are founded in scientific disciplines that are not likely to be currently represented in the socio-economic and other social science research programmes currently supported by the panel surveys. Synergies between scientific programmes and major data collection programmes in the same institution have been seen as contributing significantly to the quality of both (ESRC Strategic Review). Medical funders, such as MRC and Wellcome, approached to supplement ESRC investment in this new area, are thus likely to require assurances that scientific leadership and capacity is adequate to undertake adequately the scientific tasks involved. They will also need to be convinced that resources such as laboratories and bio-data storage facilities are readily accessible. The stored biological samples (eg blood) will then be retained for analysis after data collection in subsequent, separately funded analytes This model is used in a number of studies that collect blood samples (e.g. NSHD, NCDS, ELSA).

Options here include:

1. Consortium of individuals/organisations representing the range of required individual and institutional disciplinary strengths. (e.g. the US HRS; US Add Health; ELSA)
2. Interdisciplinary research team with new appointments as appropriate (e.g. BHPS)
3. Strong interdisciplinary scientific advisory team and panel of specialist consultants (e.g. ESRC and MRC funded birth cohort studies.)

Because of the challenges confronted by 2 and 3 our preference is for 1 with the team also backed by an advisory structure that would support the widest possible range of inputs into the design. The longer term challenge is to develop capacity in the next generation of researchers to broker research requirements between data producers and users and to undertake the interdisciplinary research itself.

5. Security and confidentiality

Organisations which have carried out fieldwork using nurse interviewers and collecting potential sensitive biomedical procedures, and laboratories which analyse the data have long established procedures for ensuring the security and confidentiality of data collected. All respondents are assigned a unique serial number and when labelling samples such as blood or saliva date of birth is used as a double check to ensure that results are matched to the correct respondent. This is particularly important when results are being fed back to respondents with potentially serious consequences. The samples are stored without names and the data files containing the results of analysis, or data such as blood pressure also contain no names but are identified using serial numbers.

There is sensitivity surrounding the possibility of identification and sample matching when biological and genetic data are collected. In this case, formal data access procedures are put in place to ensure that data are available to the academic community and bone fide researchers only. The usual procedures that are now recommended by the National Institutes of Health, US (www.fnih.org/GAIN/GAIN_home.html) (used by NCDS, ELSA, Caerphilly study in UK and Framingham Study, SWAN, Add Health and others in the US) is to provide an application procedure for the data. Applications are screened with various study specific criteria by committee and data released under license or contract. The screening procedures and license/contract approval process does require considerable resources and the processes needed to establish such procedures should be established in the early stages of the study.

Costs

Biomarker collection is expensive by the standards of social surveys because of the special field work requirements such as nurse interviewer training, instrumentation and storage, all of which add to the basic field work cost. Although it is not possible to do at this stage all the detailed work needed to cost this component of UKLHS we have made rough estimates based principally on experience in the HSE and ELSA (Appendix 2). On the assumption that the first biomarker collection would involve a home visit by a nurse interviewer and would take place in two tranches separated by a year: this gives a rounded figure of £6 million for the survey. This figure only includes the cost of the interviewing.

6. Conclusions

Britain is unusual in having a strong resource of existing longitudinal data, and a uniquely strong base of continuing longitudinal studies of individuals as vehicles to carry new questions about health related life course processes. Therefore great care should be taken in considering inclusion of health and biomarker data collections as part of the UKLHS, to be sure that such data will be collected only when the unique design features of the UKLHS, as a longitudinal study of households and families rather than individuals, are required. We believe our recommendations make best use of this and other unique features of the study and will complement this strong resource base available. Thus, in addition to recommendations which use the household design, the biomarkers are recommended from a life course and behavioural genetics perspective which utilises the wide age range of the UKLHS.

Outstanding issues include the extension of measurement in the direction of some of the new US studies, such as Add Health and NCS, that embrace fine-grained calibration of the physical environment in the home and in the locality as well as health. The extent to which personality assessment should also be included in the UKLHS will also need to be resolved.

Ultimately, the decision about whether and which data to collect must be judged in terms of the relative costs and benefits to the research and policy communities and ultimately to the public at large. We believe our recommendations make best use of the unique features of the study that will match these research needs.

REFERENCES

Baker, C. 2004. *Behavioral Genetics*, Washington, D.C.: http://www.aaas.org/spp/bgenes/publications.shtml.

Banks, J. Oldfield Z, Smith J. and Marmot M, 2006. Disease and disadvantage in the United States and in England Journal of the American Medical Association, 295: 17; 2037-2045

Barker, D. 1998 Fetal Origins of Cardiovascular and Lung Diseases. Barker DJP (Ed) RCOG Press

Burkhauser & Lillard 2005 The Case for NIA Leadership in Integrating State-of-the-Art Biomarkers into Next Generation Social-Science-Based Data Report to the National Institute on Aging.

Biddle, J. and D. Hammermesh. 1994. Beauty and the Labor Market. *American Economic Review*,

Caspi, A. et al. 2002. Role of Genotype in the Cycle of Violence in Maltreated Children. *Science* 297: 851-854.

Ding, W. Lehrer SR, Rosequist JR, and Audrain-McGovern J. 2006. The Impact of Poor Health on Education: New Evidence Using Genetic Markers, manuscript. Queens University.

Finch, C. E., J. W. Vaupel, and K. Kinsella (eds.). 2001. *Cells and Surveys – Should Biological Measures Be Included in Social Research?* Washington, D.C.: National Academic Press.

Goodman, R. (1990) The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden, *Journal of Child Psychology and Allied Disciplines,* 40, 791-799.

Knudsen, E. I., Heckman JJ, Cameron JL, and Shonkoff JP. 2006. Economic, Neurobiological and Behavioral Perspectives on Building America's Future Workforce. *IZA Discussion Paper No. 2190*. Bonn.

Kuh D and Ben Shlomo 2004 A lifecourse approach to chronic disease epidemiology. Oxford University Press

Kuh, D. Hardy R. Butterworth S. et al., 2006 Developmental origins of mid-life grip strength: findings from a birth cohort study. J. Gerontol. A Biol. Sci.Med. Sci: 61; 702-706

Lillard, D. R. and Wagner G.G.2006. The Value Added of Biomarkers in Household Panel Studies. *DIW Data Documentation* No. 14. Berlin.

Marmot M and Wilkinson R. 2006 Social determinants of health. Oxford University Press

Moffitt, T E., Caspi A, Rutter M .2006. Measured Gene-Environment Interactions in Psychopathology. *Persepectives on Psychological Science* 181, 5-27.

Rutter, M., Tizard, J.. and Whitemore, K. (1970) *Education, Health and Behaviour.* London: Longman.

Smith, J. 2004. Unraveling the SES-Health Connection, working paper IFS

APPENDIX 1

EXPERT RESPONSES

Panel of Experts consulted:

Aroon Hingorani, Reader in Cardiovascular Medicine, UCL
Tom Kirkwood, Professor in Aging, Newcastle University
Marcus Pembury, Professor in Genetics, University of Bristol
George Davey Smith, Professor in Clinical Epidemiology, University of Bristol
David Strachan, Professor in Epidemiology, St. George's Hospital, University of London
Emily Grundy, Professor Demographic Gerontology, London School of Hygiene and Tropical Medicine
Carol Dezateux, Professor in Epidemiology, Institute of Child Health, University of London
Michael Marmot, Professor in Epidemiology, UCL
Peter Jones, Professor in Psychology, University of Cambridge
Jane Wardle, Professor Clinical Psychology, UCL
Paul Pharoah, Senior Clinical Fellow, University of Cambridge,
Nick Wareham, Professor in Epidemiology University of Cambridge
Alison Stephen, Professor in Nutrition, University of Cambridge
Ingrid Schoon, Professor in Psychology, City University
John Danesh, Professor in Epidemiology and Medicine University of Cambridge
Peter Whincup, Professor in Cardiovascular Epidemiology St. George's Hospital, University of London
Chris Power, Professor in Epidemiology, Institute of Child Health, University of London
Paul Boyle, Professor in Geography, University of St. Andrews.

Responses

Aroon Hingorani (A.Hingorani@ucl.ac.uk)
To:
meena kumari (meena_ku@hotmail.co.uk)

Subject:
RE: Consultation paper on the value of including biomarkers in the proposed new UK Longitudinal Household Study

Dear Meena,

Thanks for asking me to look at this document. My comments are as follows:

It wasn't clear whether plans were in place to collect information on disease outcomes and if so, which and how. I suppose this may be in discussion

I think the case is strong for developing a DNA archive in the UKLHS, and there are reasons for thinking that the information will be complementary to other large biobanks.

1) The design is a good one for studying G*E effects
2) The existence of unrelated (spouses) and related (parents and children) individuals within the same environment offers the flexibility of studying genetic association with unrelated controls, or with family based controls, or both *within the same study* to allow good control for population stratification for any genetic association.

In relation to the list of biomarkers and measures, the preliminary list is reasonable on the basis if current evidence but the area is rapidly advancing and I think the most important thing would be to collect, if possible, sufficient quantity of serum and plasma for long term storage to increase the potential for evaluating the relevance of biomarkers for prediction and/or aetiology as new information on these accrues.

Best wishes

Aroon

From David Strachan:

Dear Meena

Here are some first reactions to the UKHLS proposal. It may be useful to have a telephone discussion if the ballpark calculations do not frighten the horses!

1. The concept of adding health-related data collection to such a large longitudinal study is obviously attractive - if it can be afforded. So, my first question is who would fund it? Evidently ESRC would be seeking co-investment from other funders. Biobank UK is testing the limits of MRC/WT/DH investment in data gathering, and I doubt if comparable levels of funding would become available there during 2006-2010. Judging from costs of the nurse examination in NCDS (£2M for 10K home visits each of 90 minutes), visits to c.40K households (with 1.5 individuals per household x 60 mins per individual) would cost c.£8M - and that is only for one visit, without any sample processing! I would allow at least £10M per sweep to include lab costs.

2. The range of "biomarkers" that you propose looks fairly standard. Since the strength of the study design is in repeated measures, the emphasis should be on measurement of change - implying multiple visits, and escalating costs. Mental health outcomes are likely to be more pliable to changing social and environmental conditions, but perhaps they will be included in the ESRC-funded interviews?

3. I am less convinced about the urgency of DNA collection. There _may_ be interesting gene*environment interactions to be studied, but these are likely to be most efficiently studied in small subgroups experiencing changing living conditions. There will inevitably be a problem of recruitment and retention of such an intensively studied household panel, and it would probably be better to keep genetics out of the initial approach, adding DNA collection at a later sweep on those families that have demonstrated loyalty to the study and have good quality longitudinal data over a number of years. I take the point about family study design, but the methodological advantages would apply mostly to two-generation families, or to sibships, whereas the focus of the social epidemiology is likely to be on spouse-pairs through middle age, retirement etc.

4. With a wide age range, we are really looking at a series of substudies (young singles, nuclear families, middle-agers, elderly, etc.) each with their own set of particular health problems. It may be difficult to justify a core protocol, except for the most obvious things, such as the basic CV risk factors, whereas age-specific problems (child health, reproductive health, disability in old age) could be neglected unless substudies are designed. Each substudy could then run into problems of power and sample size for anything but relatively common diseases (particularly if change in health status is the outcome of interest).

5. Basic information on disease prevalence, risk factor distributions, age-specific problems etc. could be gleaned from the Health Surveys for England and Scotland, more easily and more quickly than waiting for sweep 1 results.

I hope this does not sound too pessimistic. With best wishes

David

David Strachan
Professor of Epidemiology
Division of Community Health Sciences
St George's, University of London
Cranmer Terrace, London SW17 0RE

Carol Dezateux (c.dezateux@ich.ucl.ac.uk)
To:
meena kumari (meena_ku@hotmail.co.uk); j.bynner@hotmail.com; m.wadsworth@nshd.mrc.ac.uk; mblake@natcen.ac.uk

Cc:

C.Law@ich.ucl.ac.uk; Chris Power (c.power@ich.ucl.ac.uk)

Subject:
Re: Consultation paper on the value of including biomarkers in the proposed new UK Longitudinal Household Study

Dear Meena, Margaret, John and Mike

Thanks for asking for comments. My interest would be in the children and adolescents within the study and how the information - both questionnaire and biomedical - might advance the research agenda for child public health. Having said that I was not clear of the predicted age structure of the sample.

Given the timescale for responding I have made some rather rapid bullet points below and hope that we can have opportunity for further discussion and interaction at a later date. I am copying my colleagues Catherine Law, Catherine Peckham and Chris Power in but, with existing commitments, we have not had any real opportunity to discuss this jointly and so I am responsible for the comments below (and hence any inaccuracies or omissions!)

I think this study would be of epidemiological value as, for minimal extra expenditure relatively, the rich data collected can be enhanced to address important health issues that are either socially patterned or linked to factors associated with demographic economic and social change.

One point about the study design is that it would appear that it will exclude looked after children not living in a household, as well as travellers and asylum seekers. Will you be consulting children and young people about the content?

I would support enhancing the information collected with biomedical measures. I think the rationale for genetic samples and any potential impact on the other measures may be less compelling given Biobank and Generation Scotland and what we know about the necessary sample sizes for gene environment interaction. However if you decide to go ahead liaison with these groups will be essential. For children specifically blood sampling would be a big disincentive and therefore I would suggest that buccal smears would be appropriate. However given the sorts of sample sizes that would be needed and the likely spread of ages and conditions, I think the merit lies more in objective measures of phenotype and risk factors than in assessments of genotype.

Reviewing your proposed measures, obviously consideration will need to be given to blood sampling as above and many of these measures make less sense in children or else are covered within other more detailed cohorts such as ALSPAC so I would favour considering what this study could do that was not covered by these other ones. A strength of your study would be the rich economic and social (?ethnic) data and, presumably, the

heterogeneity of your sample at a household level.

Obesity and physical activity in family and environmental setting is obviously a key hot topic. The area and household influences and their relation to family income and access to types of food and opportunities for exercise locally are all recognised as important and more information on these in a household context would be helpful especially as this would allow examination of clustering of risk factors within households. Assuming that you will have anthropometry on adults in the house, for children you might like to consider different measures eg not waist circumference but bioelectrical impedance to measure fat free mass as BMI and waist circumference are recognised as having some limitations in this age group. Similarly blood pressure measurements in children are problematic - this is something the Southampton Women's Survey (Hazel Inskip) has particular recent experience with and it would be worth speaking to her. Detailed characterisation of the metabolic syndrome in children would need some thought as the blood sampling is invasive and fasting samples etc might be operationally too challenging.

There is a paucity of good information about childhood disability, so while agreeing with your four categories for health, it will be important that the questionnaires include instruments that pick up on the nature and underlying cause of a child's disability as well as the more simple questions on health state as in the GHS.

Good mental health measures will be important as this is a key issue for children and would benefit from the household measures proposed. There would be an opportunity to relate that to children's changing roles in the household (for example many children act as carers of their parents or sibs or other relatives).

Other key areas which might benefit from biomarker measures would include infections, stress, and substance (mis)use. I think these are very useful specimens and should all be sought in children.

Non invasive biological samples which are particularly suitable for children include eg oral fluids (for infection which we have used successfully in the millennium cohort study), saliva (cortisol, cotinine, drugs such as cocaine etc) or urine (cortisol, drugs), shed teeth (lead and a range of other measures).

Other biophysical measures which you should consider in all or a sub-sample would include objective measures of physical activity (in all ages) and some simple assessment of dietary behaviours. We and others are working in this area and there could be tie ins with other cross sectional surveys being sponsored by the DH and government departments.

I am sure you are considering record linkage to enhance the interviewer and respondent provided information in relation to routine health, education and social services data (as well as at a family level DWP data). I am not clear how the sample will be stratified but, if geographically heterogenous, there will be important opportunities for linking to

ecological environmental measures. For children geographies at birth as well as at current residence may also be important so that information on birth hospital and post code at conception and birth would all be highly relevant. Linkage will be greatly facilitated if you can collect NHS number.

I am sorry not be able to respond in more detail than this and hope you find this helpful. We would be interested, depending on timing and timescales, in providing input into the child elements and thinking in particular how this can link to other important sources of data on children's health, including the Millennium Cohort Study, in which I and my colleagues are very involved.

with kind regards

Carol

Professor Carol Dezateux FRCP FRCPCH FFPH FMedSci

From Ingrid Schoon:

Dear Meena

thank you for sending me the consultation paper regarding the value of including biomarkers in the proposed new UK household panel study. In principle it is a good idea to include some objective measures of individual health. My concerns are regarding the practicalities of the assessment and the overall research focus of the study. You say that education, work, retirement, income and wealth and family dydnamics are of key concern - so I think you will need a couple of questions to cover these key aspects for the study. I do not know how much time or how many visits per household will be budgeted into the proposal - yet I am concerned of overburdening the respondent households by including all the biometric measures you have included in the paper. Given a rough estimate it would take at least 20 minutes per household member to assess the measures of functioning, anthropometry and blood analysis. Given you have 3 or 4 household members, that would possibly already take up one hour of a health visitor. To my mind this might be asking too much of their time. An option to be considered is to obtain seperate funding to cover the costs for these direct assessments. A compromise would be to focus on some key assessments, such as:
rspiratory functioning (possibly using FEV)
angina
blood pressure
waist circumference only (which has shown to be highly correlated with other antroprometric measures and highly predictive, yet easier to assess - can also be done by self-assessment under supervision)
I would also suggest to use as measure for general health status the SF12 v2 (only 12 items covering physical and mental functioning)

a highly validated instrument to assess depression (such as the CES-D)
a measure of anxiety (Spielberger trait inventory)
use the CAGE
smoking
ask about drug use
add a good and short measure of cognitive functioning
and include some assessment of social values (attitudes towards family and work for example)

these would be the main areas I think need to be covered - and which could be assessed without putting too much burden on the respondent (time wise or assessment wise).

I hope you find these comments helpful

with best wishes
Ingrid

From Chris Power:

Dear Meena, Margaret, John and Mike,

Thank you for consulting me about this exciting development of the household longitudinal panel study.

I enclose brief comments below that I hope may be useful to you, although you are probably already well aware of several of these issues.

1. **The family structure** of the study is likely to be attractive to a wide range of scientific disciplines (both social and medical/biological) and is a major strength compared with other population- based surveys used for epidemiological research. The ethnic mix of the study will also be of considerable importance (especially since many other longitudinal studies are predominantly white populations, with the exception of MCS) : methods that might be introduced to increase participation over time (or capture data through other mechanisms) could be a strength of the study if this were built in as a methodological component.

2. **Inclusion of biomarkers** (possibly including genotypes) is likely to be of value to health economists interested for example, in improved control for health status. Some existing studies that potentially might be used for this purpose (eg the 1958 birth cohort) have consent only for medical research and not for a broader range of purposes. With appropriate consents, the proposed new household longitudinal panel study could fill an important gap. However, leaving aside any ethical issues for example relating to blood sampling of children, justification of the costs of assays /genotypes on this large sample might be difficult given other resources that are likely to be widely available to the research community, such as ALSPAC and Biobank.

3. **Study aims**: Good measurement of health is essential for the study to have value for epidemiologists and well deserves the prominence given in your outline objectives. For

epidemiologists to actively engage in the study there is clearly a need to develop a strategy which deals with the problems inherent in multiple purpose studies in which data quality is compromised because of the need to (i) reduce respondent burden and (ii) ensure that different topic areas are covered. Enthusiasm from epidemiologists might therefore be enhanced if the feasibility of collecting good quality data (for example health related habits, such as diet and physical activity) within the context of the new study could be established at an early stage.

4. I would be happy to contribute to a discussion of details of specific biomarkers, particularly those with which I have recent experience in the biomedical study of the 1958 cohort. Your recommended list looks comprehensive, although how the measures apply across the different age groups (from the youngest children in the family to the oldest adults) will require consideration?

The study could clearly be of considerable interest to epidemiologists and I will be interested to see how this develops.

With best wishes
Chris

Precis by M. Kumari of telephone conversation with Emily Grundy:

1. What is the source of funding given the other cohort studies?
2. Will the measures be different by age group? Thus, we propose hand grip strength to assess functioning but ELSA uses a wider array of measures, would we supplement at different life stages?

APPENDIX 2
COSTS

ESTIMATED COSTS FOR NURSE WORK ON PROPOSED UKHLS
28/9/2006

Assumptions

Inflation: figures not inflated up to the year when fieldwork would take place. They assume fieldwork in 2007.

Sample:

- 40,000 addresses are divided into assignments of about 20 addresses distributed across 12 months of the year (nurse work would be divided into two so only half the households households contacted in each wave – half the sample points contacted in each wave for nurse visits).
- Nurse visits at 34,000 addresses (85% response to nurse visit assumed with a base of households responding at wave 1)
- Assume 180 nurses work on project and need nurse equipment (assuming only half households are contacted at wave 3 and half at wave 4)
- Assume 400 interviewers need height and weight equipment

Length of nurse visit:

Adult (16+)	30 mins
Child 11-15	20 mins
Child 4-10	15 mins
Child 0-3	5 mins

Content of nurse visit:

- **Blood samples: (11+) (Costs of analytes and blood storage not included)**
- *Grip strength (16+)*
- *Blood pressure (5+)*
- *Lung Function (7+)*
- *Infant length (0-1)*
- *Demi-span (65+)*
- *Waist and hip (11+)*
- *Body fat (not included in equipment costs) (7+)*

- *Height and weight (2+, 0+) - in interview so fieldwork time not costed. Adds 10 minutes to interview time for 2 adults.*

Estimated costs

- Nurse interview for 100 households (c. 167 adults and 42 children): £14,300
- Disposable equipment for 100 households: £1,200
    - So for each of wave 3 and 4 (covering 17,000 each): £2,635,000

- Briefing 180 nurses (each wave with a nurse visit): £41,000
- Other costs (print, postage, accommodation, travel, recruitment) (each wave with a nurse visit): £120,000
    - So for each of wave 3 and 4 (£161,000)

- Fixed equipment – one off purchase for project: £310,000

TOTAL FOR INITIAL NURSE VISIT (AT WAVES 3 AND 4): £5,902,000

This does **not** include:

- Laboratory analysis
- Laboratory storage
- Cost of health questions in interviewer visit
- Research costs of designing and implementing nurse visit
- Inflation beyond mid 2007