Enabling transdisciplinary research in the UKHLS: Incorporating biomarkers and pathways into research on the interplays among social, economic, behavioural and health sciences

John Hobcraft

University of York
Executive Summary

The UK Household Longitudinal Study (UKHLS) is intended to be a laboratory which will continually be at the frontiers of research on social, demographic, behavioural and health sciences. In achieving this goal it is essential that the study engages with the rapid advances in the biological and life sciences, in order to enable genuine advances in transdisciplinary research.

The UKHLS has several advantages for exploring these interplays: it is prospective; it covers of the full age-range and all household members; and interviews will be annual. The target sample sizes are large: around 40K households and 100K individuals; about 13K children aged under 16, compared with from 5K to 18K in the birth cohort studies; around 35K aged 40-69, compared with 25K of the 500K Biobank sample for whom repeated interviews are likely; and around 15K aged 65 and over, compared with 5.6K in ELSA.

Two core elements, biomarkers and stable characteristics (sometimes referred to as intermediate phenotypes), are proposed and justified in some detail for the UKHLS, each requiring about five minutes of ‘interview’ time each year, but with changing content. Over time, this permits an accumulation of information on biomarkers and on key stable but non-biological markers that help to shape behaviour and play important roles in linking biology to behaviours or outcomes.

The first key element is collection of biomarker information by minimally-invasive methods that can be carried out by survey interviewers with appropriate training. The roster of such biomarkers could include: anthropometry; saliva for DNA and subsequent genotyping; dried bloodspots for subsequent assay; physical functioning measures; and perhaps in due course suitable body fluids (blood or urine) for Nuclear Magnetic Resonance study of metabonomics.

The second key element is to collect information on relatively stable characteristics that nevertheless shape behaviours, especially those that might mediate or moderate genetic or other biomarker related susceptibilities for behaviours or outcomes of interest. Examples include: cognitive ability/ functioning; personality traits; cognitive styles – decisiveness, impulse control, trust, etc; mental schemas or heuristics on economic or family behaviours; motivation, self-esteem, etc. These measures would be collected in a variety of innovative and appropriate ways that could include: traditional questions; use of vignettes; games drawn from behaviour economics; computer-based approaches, etc. Because they are relatively stable over time they would be collected infrequently and rotated as appropriate.

The combination of a rich set of biomarkers and of intermediate phenotypes (or stable characteristics) would enhance the value of the UKHLS for studying gene-environment interplays, for exploring pathways from genes through stable characteristics to social, demographic, economic, psychological and health outcomes and behaviours, and for illuminating the interplays among these disciplinary domains. Examples are provided.

It is proposed that an integrated package of biomarkers (and key related indicators) be developed to cover a three to five year period and funding sought, with help from
the ESRC, from a consortium that could include Wellcome Trust, MRC and BBSRC. Some indicative costings are also provided. Moreover some possible mechanisms for enabling the requisite transdisciplinary research to be brought to fruition are outlined. The importance of this research agenda for science and for policy and the unique investment in the UKHLS require a substantial and sustained investment to bridge the social, health and biological sciences.
Introduction

The UK Household Longitudinal Study (UKHLS) is set to become a flagship project for UK social, demographic, economic and behavioural science which can also have a profound impact on health research. Its scope and ambition is unprecedented for a prospective household survey, covering 40,000 households and an estimated 100,000 individuals, with substantial annual interviews. It will thus become a storehouse of ‘phenotypic’ information on outcomes, behaviours, and their determinants, contexts, and consequences for a sample of unprecedented size. Moreover, the focus on households will enable the exploration of the ways in which individuals’ behaviours and life-courses are shaped by their co-residents, especially family ties across and between generations.

From the outset, the need for the UKHLS to engage with biological and health sciences has been recognised (e.g. Kumari et al 2006) and it is increasingly apparent that understanding of economic, social, demographic and behavioural science requires attention to the interplays of nature and nurture (e.g. Rutter 2006; Hobcraft 2006). For the UKHLS to remain at the frontiers there will have to be a continuing interplay of social scientists with geneticists, neuroscientists, microbiologists, epidemiologists, etc to enhance understanding, learn from animal research, and explore new interdisciplinary hypotheses. This paper addresses some of the issues and possible mechanisms for taking such research forward through the UKHLS.

Some aspects of engagement with the biological and health sciences are inherently expensive and will require additional funding (e.g. genotyping, other assays, any possible brain imaging, nurse visits, etc). However, the potential of such a large sample for examining these interplays is huge: much genetic research is hindered by lack of power and the demand continues to grow for ever larger samples. One major intrinsic advantage of prospective studies is their potential for exploring gene-environment interactions (Manolio et al 2006; Rutter et al 2006); however some of the most important findings from the Dunedin Study have not required very large samples, but rather careful informed candidate gene research (Moffitt et al 2005 & 2006). We stress that incorporation of the existing BHPS sample would permit early exploration of these interplays if suitable genotyping were done.

The potential value of the UKHLS for linking biomarkers, health and the social sciences needs to be placed in the context of existing or newly established prospective studies. It is unusual in three respects: sample size, aiming to cover all ages, and to cover all members of the household. The UK has a widely envied series of birth cohort studies, which began in 1946, 1958, 1970, 1990/1, and 2000/1. Biomarkers, including DNA, have been collected in the 1946 NSHD cohort and in the 1990/1 ALSPAC cohort and in a recent round of the 1958 NCDS cohort. In none of these studies are the linked data yet accessible in the public domain. Effective samples are typically around 8/10K individuals, though only 3K for the 1946 cohort. All of these birth cohort studies are nevertheless of huge potential value for the study of the interplays of biomarkers, health, and social science outcomes and several have the advantage of already being long-running studies. The UKHLS will have an initial sample of around 13K under age 16. Much newer is ELSA with an initial achieved sample of 12K (mainly 50 and over, of whom 5.6K are 65 and over); the UKHLS should have an initial sample of about 15K aged 65 and over.
The UK Biobank aims to recruit 500K individuals aged 40-69 over the next few years, but currently only plans to collect limited socioeconomic and lifestyle information at entry; there are then plans to get ‘representative’ updates every 2/3 years for samples of about 25K (UK Biobank 2006); in contrast, the UKHLS should have around 35K initial recruits aged 40-69 and will monitor changes over time in their circumstances and health on an annual or periodic basis. Thus the UKHLS is likely to have at least as large an effective size as the Biobank for exploring gene-environment interplays over time, where it has to be the environments that change and not the genes (and the study of changes in gene expression would require new biological samples in either study); moreover UKHLS will obtain repeated measures on the same individuals over time, rather than on changing random selections from a much larger population.

Thus, the UKHLS has sample sizes that are typically as large as or larger than many existing prospective population-based studies and also offers the opportunity to explore within-family associations. Of course, being a general purpose study of a large sample, the UKHLS may not be able to collect such rich information as some of the more in-depth studies (e.g. ELSA or NSHD on ageing) but, as is discussed below, the opportunity to collect modest amounts of information on biomarkers and ‘stable characteristics’ annually means that comparable richness may be achievable over time and the larger sample sizes should make for greater opportunity for gene-environment studies. It is perhaps most salutary to realise that the sample size is of the same effective order for some purposes as the apparently much larger Biobank and the likely richness of information will make the UKHLS an important data source, provided the outcomes being studied are fairly prevalent and data linkage to health and other relevant administrative records is achieved.

The underlying philosophy adopted here is that the emphasis in the UKHLS has to shift significantly towards understanding and away from description. Intrinsic to a decision to carry out a prospective study is that continuity and change over the life course matter and that understanding is far more likely to emerge through examination of pathways and interplays over time. For example, detailed income dynamics may be best studied using administrative records; and retrospective event histories cannot help understanding of processes very much. Tough judgements have to be made over these and similar trade-offs and a key guiding principle should be the contribution to scientific knowledge, since it is ultimately better understanding that leads to truly informed evidence-based policy. The UKHLS, as a flagship project, can and must aim to meet the highest standards of innovation in content and approach and to make a major contribution to the advancement of the social and health sciences. Ensuring policy relevance, by tackling important questions, is also of critical importance.

In order to achieve such an advance there is a need to consider a profound reorientation in the approach to the social sciences (including social, demographic, economic and behavioural elements) which builds upon some of the best features of psychology and health sciences in interlinking ‘alleles, brains, and contexts’ into the study of development (of careers, relationships, cognition, personality, and health) across the life-course. Behavioural economics has seen a profound engagement of
psychology and economics, which has radically altered the way in which (some) economics is done and is now beginning to reach back to genetics and neuroscience.

Most major large-scale prospective studies in the US are now collecting biomarkers, usually including saliva (now the preferred vehicle for genotyping, but also for cortisol) and dried blood-spots (see McDade et al forthcoming for a list of well-validated assays): for example AddHealth, Health and Retirement Study, Fragile Families Study, LA Family and Neighbourhoods Study, and the Wisconsin Longitudinal Study; the massive National Children Study (of a cohort of 100,000) will also collect very extensive biomarker information. A group at Harvard (led by David Laibson) are currently exploring a large panel of genetic markers and links to economic behaviour in an Icelandic population. At this stage most of these studies are adopting a candidate gene approach, often involving the most commonly identified markers on fairly well-documented genes, but sometimes also exploring combinations of markers within these genes and increasingly considering less thoroughly explored candidate genes. Equally there is a fairly common list of analytes currently being assayed from the dried blood spots. The UKHLS can learn from and build on this experience.

However, after extensive discussions with Principal Investigators and funders of these studies I have become convinced that more is needed. In particular, more attention needs to be paid to pathways from genes through the brain and through ‘relatively stable characteristics’ to behaviours and outcomes. This theme has partially emerged in the psychological literature on intermediate phenotypes (sometimes called endophenotypes) in discussion of the search for gene-environment interactions (see Hernandez and Blazer 2006, pp 72-82), but should be construed much more broadly. Many social scientists have become more aware of (relatively stable) personality traits and inclusion in surveys of some version of the NEO five-factor scale is now often advocated – this is an example of a feature of the individual that changes slowly over the life-course, but is related to much downstream behaviour (though links back to specific genes have proven somewhat elusive - see Ebstein 2006 for a comprehensive review); cognitive performance is another such powerful shaper of downstream outcomes (with similar problems in identifying genetic markers - see Plomin et al 2006; Plomin and Spinath 2004). However, there are many other relatively stable attributes of individuals that probably need to receive much greater emphasis in exploring the pathways to behaviours and outcomes: commitment, trust, pair and parent-child bonding, self-esteem, risk aversion, impulse control, etc. At a slightly more abstract, but nevertheless important, level might be tapping into heuristics or schemas that people use in everyday (or key) economic or family decision-making. There is a case for regarding depressive tendencies and life-satisfaction as being relatively stable characteristics, as may be some attitudes, preferences and beliefs. Moreover, we advocate that a variety of appropriate measurement approaches be used, including for example trust games (Ermisch and Gambetta 2006), vignettes, or computerised assessments such as implicit association tests for assessing gender bias or racial prejudice, etc, in addition to more traditional interview questions. Some of the more novel approaches should be tested in the innovation panel.

It may seem perverse to be recommending a significant emphasis on relatively stable attributes in a prospective study, where the emphasis is often on change. However, the annual nature of the UKHLS means that a significant battery of information on these
fairly stable characteristics can be built up over time. This can then enable a much richer exploration of some of the key pathways that almost certainly influence a variety of more transient behaviours and outcomes. These life-course experiences (e.g. parenthood, divorce, unemployment) may then play an important part in feedback to bring about changes in the relatively stable characteristics measured a few years later. For example, Kiernan (1986) showed that those who divorced had higher neuroticism scores before marriage and not just after divorce and Easterlin (2006) showed a similar relationship for before and after low life satisfaction scores for those who divorce. Exploration of gene-environment interactions is likely to be most powerful where there are before and after measures of the outcome (e.g. the seminal study by Caspi et al 2003 regarding the interplay of life stress with genetic susceptibility in the onset of depression).

Biomarkers

There are three main approaches to the collection of biomarkers: getting respondents to attend clinics, having nurses (or other medically trained personnel) visit respondents in their homes, and using minimally-invasive approaches with trained survey interviewers (sometimes supplemented by mail-back items). Evidently these three approaches progressively decrease in cost and in the complexity of the measurements that are possible. Given the size and scope of the UKHLS, the additional costs of nurse or clinic approaches would be substantial (Kumari et al 2006 provided an estimated additional survey cost for a nurse visit in the UKHLS of about £6 million). However, the collection of useful information on biomarkers using survey interviewers has progressed significantly over recent years, with development of lower cost or more portable technologies and of increasingly sophisticated measurement possibilities through these minimally-invasive approaches (e.g. Lindau and McDade forthcoming).

The annual sweeps of the UKHLS make it possible to collect different minimally-invasive biomarkers in successive waves and thus to assemble a rich collection over time, whilst only allocating a few minutes (target five minutes) in each wave. Thus we would envisage consideration being given to the collection of: saliva, now the preferred minimally-invasive approach for obtaining DNA; dried blood spots, where the number of reliable assays is rapidly increasing (Mei et al 2001; McDade et al forthcoming); anthropometric measures of height, weight, waist circumference, etc and possibly bioimpedance; spirometry to assess lung function; blood pressure and pulse rate; a lipid panel; and other possible measures, such as a four-sample mail-back regime of saliva samples to obtain reliable indications of cortisol (and other assays), or more speculatively for the future urine samples for analysis of metabolites using NMR or the use of portable brain-scanning devices.

Each such biomarker would be subject to separate consents. Protocols for some would need testing in the innovation panel to devise ways of minimising non-response and to ensure it was clear to respondents that refusal to provide such biomarkers did not prejudice their inclusion in the Panel.
One of the key issues concerning biomarkers in the UKHLS is to identify those which are likely to add to our understanding of social, economic, and demographic behaviours, where linkages to biological pathways are as yet less well established than for linkages to physical or mental health. Achieving an appropriate balance and choosing additional promising biomarkers beyond the health domain will be one of the major scientific challenges facing the UKHLS. In the meantime, a case can perhaps be made for giving priority to genotyping and to a wide variety of assays from blood spots, where the results open greater possibilities for exploration of linkages to behaviours and outcomes that are central to the social sciences, whilst providing measures that are also of established value to the health and psychological domains. This is not to underestimate the value and importance of incorporating more focussed health measures into the biomarkers and permitting exploration of the biosocial pathways involved in understanding health processes.

Moreover, whenever funding and collection protocols in the health domain are considered for the UKHLS it is essential that health is interpreted as ‘a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity’, in line with the WHO definition. A key priority for the UKHLS will be to ensure that such measures are available for research on a wide range of social, economic, behavioural and health domains and not restricted by narrow protocols (as for example happened for the genotyping in the NCDS).

We now turn to a more detailed consideration of some of the specific issues involved in the proposed biomarkers for consideration for the UKHLS.

Salivary DNA

Although approaches may develop further, there is currently a remarkable convergence on saliva as the preferred minimally-invasive approach to collecting samples to enable genotyping. Buccal swabs and mouth wash techniques have proven problematic because the yields of DNA are usually quite low (as with blood spots) and several common candidate genetic markers (such as those within DAT1 and DRD4, dopamine pathways linked to ADHD for example) involve identifying VNTRs (variable number tandem repeats) which demand more DNA for genotyping than the more common SNPs (single-nucleotide polymorphisms). Most large-scale studies in the US (especially those with recent or upcoming fieldwork) and several worldwide (e.g. a 40000 person cohort at the Karolinska Institute) have opted for a single approach, using the Oragene product to collect and stabilise saliva samples. Saliva collection is possible by professional trained interviewers (or even through mail-out/ mail-back, as has been successfully used in the Wisconsin Longitudinal Study), whereas the only realistic alternative of collecting whole blood requires a phlebotomist or trained nurse; the saliva specimens are stable at room temperature for quite long periods and can be mailed; and DNA purification is simple (with a typical yield of 80 to 110 micrograms of DNA) and the resulting material compatible with major genotyping technologies and gives reliable results. These are a formidable list of advantages for a large-scale household study. I am told the product costs around £5 per respondent. Thus the collection of DNA in the UKHLS is technically feasible and might cost up to £500K for equipment.
However, such collection is not justifiable without clear plans as to what the benefits of genotyping might be and some evidence that funding for this would be forthcoming. Provided the protocols enable it, DNA can be stored in laboratory grade freezers and used for further genotyping as new methods and hypotheses become available or as technologies become ever cheaper. There are two approaches to genotyping that need to be considered: candidate gene approaches and genome-wide scans.

Genome-wide scans require large samples to provide sufficient power, especially for detecting gene-environment interactions (e.g. UK Biobank 2006, p8) and can benefit from having observations on several family-members (e.g. mother and father and offspring) to permit linkage disequilibrium analysis: the UKHLS clearly has many benefits in this respect. The UKHLS also has potential benefits for genotyping studies because of the strategic oversampling of ethnic and regional populations, allowing a sharper focus on population stratification issues.

One strategy worth consideration would be to carry out a genome-wide scan on the current BHPS sample, where we already have 15 years of observations on the phenotype and then genotype the new sample subsequently (probably after gaining a few years’ responses) to test for replication and new hypotheses. A full genome-scan using the latest Affymetrix 5.0 or 6.0 microarrays (with around 900K markers identified) on 15K BHPS respondents would cost two to three million pounds for the arrays and assay kits; less detailed microarrays would be cheaper. The scope for exploratory work examining indications of genetic markers being associated with particular outcomes or measures as main effects would be enormous and some searching for gene-environment interactions may also prove possible. A project of this kind would involve extensive collaboration with geneticists and statistical geneticists and might, for example, be developed with the Sanger laboratories or the Wellcome Trust Case Control Consortium (2007). Funding might then be sought from the Wellcome Trust.

The alternative approach is to identify candidate genes: ones that have a sound basis for being likely to be linked to the outcome or behaviour, either from the existing literature or in animal studies, and only genotype for already identified markers on these genes. There is much interest in identifying gene-environment interactions (Rutter 2006) for behaviours or other outcomes, although these have proven somewhat difficult to identify and establish (see Rutter et al 2006, Moffitt et al 2005 and 2006 for strategic discussion). One of the best replicated gene-environment interactions is that linking long/short repeat markers on the serotonin transporter gene in combination with experience of life stress to the onset of depression and related outcomes (Caspi et al 2003). This finding is remarkable in coming from a relatively small study (the Dunedin Study, with only around one thousand respondents) and indicating the value of careful assessment of likely pathways and measures. Caspi et al (2002) have also provided evidence of a gene-environment interaction involving the interplay of a marker on the MAOA gene with child abuse in altering the risks of adult anti-social behaviour. There have been recent indications of an interaction of a ‘haplotype’ (a combination of two different markers) on the dopamine transporter gene interacting with maternal alcohol consumption (and possibly cigarette smoking) during pregnancy to alter risks of ADHD (Brookes et 2006).
There is a much larger literature that explores the ‘main effects’ of genetic markers for a wide range of outcomes. The resulting effects are often small and elusive and it is likely that gene-environment interactions will prove more important, though harder to find. Gene-environment interactions involve the interplay of nature and nurture and are a more plausible mechanism for understanding change or the onset of behaviours, disease, or other outcomes. However, one of the reasons that gene-environment interactions have proven difficult to find is that they probably involve epigenetic effects that alter the expression of the gene in lasting ways (e.g. through methylation) rather than simple variations in DNA markers. Within the life of the UKHLS there will be huge progress in this area, with the likely result that a single genetic sample will prove inadequate - although the DNA itself does not change over time, the gene expression does. I understand that it is possible to study methylation (a key element of epigenetic changes) from salivary samples (personal communication from Dan Notterman), but there are still a host of issues requiring development and understanding: much gene expression is tissue-specific and many of the genes of interest are those coding for neurotransmitter pathways, so would an analysis of salivary or blood DNA samples show important changes? There are also likely to be major advances in knowledge, understanding and measurement of RNA over the next few years. However, such developments suggest that early collection and storage of DNA samples for the whole UKHLS sample may be desirable in order to permit the possibility of studying gene-expression changes linked to experiences, as has been a strategy for data collection in ALSPAC (personal communication from Marcus Pembrey).

Using a candidate gene approach would require a careful consideration of which genes to genotype and which markers to identify within those genes. In terms of human behaviours most effort so far has gone into study of the serotonin and dopamine systems and it is almost certain that any panel of markers would include some of these. Ever more markers related to health outcomes are being identified, such that, for example, no serious study of dementia could now be undertaken without genotyping on the APOE epsilon 4 variant. However, Walley et al (2006) list 22 genes that have shown five or more positive associations with obesity. Thus a list of candidate genes could get quite large and judicious selection will be required. Judgements will have to be made as to whether to just use the ‘usual suspect’ markers within a gene or to carry out some sequencing and at what stage a marker or gene is worth investigation.

An example of this issue is given by the case for genotyping on the ‘bonding’ genes for oxytocin and vasopressin receptors (see Hobcraft 2006 for an extended discussion and many references). There is substantial evidence from animal studies for such linkages in pair-bonding; some indications from fMRI scans of humans that suggest oxytocin and vasopressin receptors play a role in a mother responding to her child or in responses to romantic partners; we know oxytocin is a key hormone in breastfeeding; oxytocin is released during the human female orgasm, as is vasopressin for the male; and recent experimental economic studies have shown that intranasal oxytocin sprays enhance trust. So there is a reasonable chance that these neurotransmitter pathways are involved in demographic or pair-bonding behaviours and in economic trust and thus may be of real importance for social science. But at what point is such evidence deemed strong enough to explore in UKHLS?
Indicative costs for genotyping come from K bioscience, who would charge £362K for 50 SNPs on 100,000 samples (or about the same for 100 SNPs on 50,000 samples). Some key markers of interest are not SNPs, but for example VNTRs, which may cost more to genotype. I would strongly recommend genotyping a panel of markers for at least the existing BHPS sample as soon as possible, so as to enable social scientists to begin exploring these important issues linking nature and nurture.

Dried Blood Spots

Collection of dried whole blood on filter paper is far less invasive than venipuncture and advances in immunoassay technology mean that many analytes can be identified with much smaller volumes of blood; there are also rapid developments in simultaneous assessment of multiple analytes from a single sample. Typical protocols in large-scale surveys involve finger cleaning with isopropyl alcohol, use of a sterile disposable lancet (as used by diabetics), collection of five blood spots of standard diameter (approximately 50µL per spot), and subsequent drying and storage. Once dried the samples can be mailed to a laboratory and kept frozen with desiccant; however there is a growing body of evidence that most analytes are stable for at least two weeks at room temperature. Moreover, samples are less vulnerable to degradation through freeze/thaw cycles than whole blood (see Mei et al 2001 and McDade et al forthcoming for a much fuller account).

When collecting blood spots in the field it is feasible to add collection of a drop of whole blood on a finger stick that can then be used in an onsite digital auto-analyzer to obtain a lipid panel analysis of total cholesterol, high density lipoprotein cholesterol and total triglycerides, as is being done in the AddHealth, HRS, and LA FANS studies in the US, though this significantly increases cost.

McDade et al (forthcoming) provide a list of 45 analytes judged relevant for population studies for which reliable assays have already been developed (Mei et al 2001 provide a longer list, many of which are too rare to warrant inclusion in UKHLS). Most assays can be carried out using a standard hole-punch, with seven punches being typical from one blood spot. Several assays are being used in major population studies, including: HbA1c (glycosylated haemoglobin), a measure of glucose intolerance, which has been effectively used in contrasts of health in ELSA and the US HRS (Banks et al 2006); high-sensitivity C-Reactive Protein, an inflammatory marker implicated in cardiovascular disease; Epstein-Barr Virus antibodies, an immunological marker of chronic stress; and DHEA-S where high levels appear to be implicated as prosurvival markers and also related to cognitive performance. While assays exist for cortisol, widely used as a stress marker, the diurnal variation in circulating levels makes this an unlikely candidate for inclusion in a panel of analytes.

During the life of the UKHLS, there will almost certainly be rapid progress in the analytes identifiable from dried blood spots and widespread scientific consultation on appropriate measures is needed. There are currently potential problems in finding laboratories that can handle dried blood spot analysis on a very large scale and careful planning and piloting is needed. However, this approach seems well worth considering for inclusion in the BHPS since there are likely to be major returns. The supplies required in the field cost about £1 per respondent and the laboratory analysis
perhaps in the range £2.50 to £10 per respondent including labour and materials, depending upon the assays required. Adding an in-field lipid test might cost an additional £4 per respondent for a lipid panel test strip and perhaps £250 to £400 per interviewer for a good quality professional hand-held auto-analyzer (with printer).

Other biomarkers

There are a wide variety of other biological or physical measures that can be considered for inclusion in the UKHLS. It seems desirable to restrict such measures to those that meet several key criteria:

- they must have a sufficiently high prevalence in the general population;
- they should be relevant across a broad age-range;
- they should be capable of minimally-invasive measurement in the field and take a maximum of about five minutes of respondent time;
- they should be of relevance to as broad a range of studies as feasible, so that preference is given to those that can inform research across a range of disciplines or domains;
- however, there also needs to be at least one very clear hypothesis for which the biomarker is critically relevant, whether as an outcome or a pathway.

Anthropometry is an important domain for collection of physical information, especially given the policy interests in obesity and its consequences, and is a likely high priority for the UKHLS. At a minimum, height and weight should be measured, but there is increasing evidence to suggest that waist circumference (possibly supplemented by hip circumference, although interviewers often find this embarrassing to do themselves) is a key risk indicator; there are also improving possibilities for measuring bioimpedance in the field, using modified scales (as is being considered for the MCS) although it is not clear that such portable technology is reliable yet. It seems unlikely that UKHLS will collect very detailed information on diet, thus limiting the study of obesity per se.

However, there are possibilities for studying levels physical activity using small accelerometry monitors over a week or more, as is being explored for the MCS. This would provide much detailed information relevant to the study of health and to obesity.

Collection of urine in field surveys is feasible, although diurnal fluctuations in hormonal levels and neuroendocrine metabolites complicate measurement from these samples. Urine needs to be refrigerated within two hours (or at least cold packaged). But there are still a wide variety of assays that could prove valuable in elucidating pathways to health and other outcomes. Some experiments are under way that involve collection through filter papers or diapers; although validated assays are still being developed; these approaches may reduce cold-chain requirements (Lindau and McDade forthcoming). Recently there has been pioneering work examining metabonomics in large population studies using Nuclear Magnetic Resonance techniques to explore statistical patterns in relation to disease (Singer et al 2007). As such studies develop and clearer signatures of predisposition to or incidence of disease are identified, along with significant cost reductions, useful possibilities may
occur for incorporating this approach into a very large population study like the UKHLS.

Portable assessment of blood pressure and pulse-rate is well-established technology usable in the field, preferably with automated readout or storage of the results. This seems a strong candidate for inclusion in the UKHLS.

Measurement of lung function in the field is now possible using portable spirometers (e.g. the Fragile Families Study and LA FANS) and can enable study of childhood and adult asthma and reductions in lung function with ageing.

A wide range of other innovative approaches to measuring biomarkers are being tried or considered in population surveys, including ambulatory electrocardiograms (expensive), portable audiometers, and self-administered vaginal swabs (successfully used in NSHAP to examine HPV infection); and attempts are being made to develop truly portable MRI to enable brain scans in the field (Lindau and McDade forthcoming). The UKHLS will need to keep a watching brief on such developments and assess which are likely to provide adequate cost-benefit returns to knowledge if implemented in a large-scale population study.

Collection of biomarkers adds complexity and expense to population studies, raising issues of non-response, interviewer and respondent fatigue, logistics and cost. But biomarkers need integrating into large population studies, especially prospective ones, to enable better understanding of pathways and processes and the interplays between biology and behaviour over the life-course. Decisions about these trade-offs need to be taken through a careful evaluation of the benefits and are more likely to be favourable (and funded) if there are clear hypotheses and supporting evidence to suggest a real return to science (e.g. Ewbank forthcoming).

### Stable characteristics

Just as there is a danger that biomarkers are too dominated by health, to the exclusion of social, demographic, economic or psychological concerns, there is some risk that measurement of what we term here ‘stable characteristics’ becomes overly dominated by psychological attributes. Both risks of such overspecialisation are understandable, since health and psychology respectively have progressed far further in these domains of biomarkers and traits. Moreover, since one of the key reasons for emphasising the need to focus on relatively stable attributes is their putative or known roles in pathways to behaviours, it is likely that mind and brain will feature heavily. We again emphasise the need for imagination and innovation to ensure that the measures reflect as wide a range of stable characteristics as can be justified in the social, economic, demographic, health and behavioural domains. Much of the required development work should engage social scientists with psychology, as has happened to such striking effect in behavioural economics.

Our starting point is that these fairly stable attributes can serve both as important mediating or shaping characteristics in the pathways to often diverse behavioural outcomes and as indicators or outcomes in their own right, where change is quite noteworthy because it probably required significant experiences or stress and may
have major downstream consequences. Moreover, such stable attributes can also play a role in understanding feedback mechanisms that elicit altered biological responses. We stress that the characteristics are only relatively stable and thus subject to some change, requiring periodic but far from annual monitoring. This enables collection of a fairly wide range of such attributes over time, even assuming that around five minutes of interview time is available each year (though addition of mail-back self-completion questionnaires might be considered).

We envisage use of a variety of collection approaches for gaining insight into such stable attributes and anticipate that this would enhance respondent cooperation by adding variety to the interviews: for example vignettes, experiments or games, or implicit association tests on prejudice could be added to traditional questionnaire approaches. Moreover, some thought should be given to using multiple respondents (where available e.g. in the household or teachers for children) to answer fewer questions on themselves and similar questions on others, permitting use of latent construct and structural equation modelling approaches to improve measurement.

Examples of relatively stable characteristics from different domains that might be considered, in some cases requiring methodological development in the innovation panel, are:

- **personality**: the five-factor or big five personality traits, self-esteem, self-efficacy
- **cognition**: cognitive performance, cognitive styles such as impulse control or decisiveness
- **economic**: trust, risk-aversion, impulse control, key decision making heuristics, motivation, well-being, soft skills
- **social**: support networks, world views, social capital, cultural schemas, gender roles, ethnic prejudice, altruism, well-being
- **family**: parenting styles, bonding experiences, commitment, trust, altruism, heuristics or schemas on managing relationship stress or teenage pregnancy
- **health**: exercise behaviour, smoking, drinking, diet, mental health, well-being
- **emotional**: affect/bonding, depressive tendencies, anger management, coping with stress, well-being
- **preferences, beliefs, values, and attitudes**: many of the above

The list is formidable and quite deliberately repeats some items under more than one heading to indicate the importance of thinking across domains. Establishing priorities and dealing with competing demands remains a challenge, but these measures are generally highly likely to:

- be relatively stable over time
- be of real importance if they show significant change
- matter across a broad age range
- play a key role in understanding pathways to behaviours or outcomes
- to interplay (e.g. across domains) in pathways from one to another or in shaping yet other outcomes

Some examples of the proven or likely importance of these measures in pathways are:

- the interaction of the serotonin transporter 5-HTT LPR marker with experience of life stress in altering depressive status
• possible links of the oxytocin receptor gene through trust to economic or family behaviour
• linking behaviours and these characteristics to allostatic load as measured by biomarkers and perhaps on to health outcomes
• exploring the links of vasopressin or oxytocin receptors, bonding characteristics, relationship stress schemas, prior neuroticism and life satisfaction to partnership formation and breakdown

Thus, these relatively stable characteristics are an integral and essential part of the ambitions to incorporate biomarkers into the UKHLS and to improving our understanding of behaviour and outcomes across all of the domains of interest, including health, social, economic, demographic and psychological sciences.

Conclusion

Much detail on the timing and frequency of collection both for the biomarker measures and the ‘stable attributes’ remains to be sorted out, as does the careful evaluation of which measures to include. A case can be made that some biomarkers and stable attributes are more likely to change during childhood or old age, thus leading to pressures for age-dependent frequency of observation (as argued by Kumari et al 2006 in the context of health aspects of the UKHLS). It would be complementary if other elements required measurement more frequently during early to mid-adulthood: economic and family related ‘stable attributes’ may fit this criterion; as may the event-triggered components of the UKHLS.

For biomarkers I would advocate salivary DNA, anthropometry, and dried bloodspots as the initial priorities, since they are likely to contribute most widely to the broad scientific goals of the UKHLS. This suggestion would involve genotyping a panel of possibly 100 to 200 carefully chosen candidate markers for the entire sample and seeking to carry out a genome-wide scan for the continuing members of the existing BHPS sample, for whom up to sixteen years’ phenotypic information has already been collected. I place this first (although there is a good argument for waiting until at least the second wave for new sample members) since this is such a fast moving field. It is really important to get social scientists, who know a great deal about the non-genetic aspects, collaboratively engaged in such work, since ‘the genetics of behaviour is much too important to be left to geneticists’ (Plomin 2001). It would be desirable to include bioimpedance measures in the anthropometry, though costs and logistics may constrain this. Consideration should be given to adding a lipid panel measurement to the dried bloodspot collection round, provided the additional funding can be found. The panel of assays for the dried bloodspots would need to be planned carefully. It would also be desirable to get blood pressure and pulse readings fairly early in the study. It is important that the protocols enable storage of salivary DNA and dried bloodspots for future analysis and are worded so as to enable use for the broad range of goals of the UKHLS. Other biomarkers and repeat measures would follow as appropriate.

One possible approach to obtaining funding for the collection and analysis of biomarkers would be to develop an integrated and prioritised package of minimally invasive biomarkers for inclusion in the UKHLS over a three to five year period and,
with help from the ESRC, seek consortium funding by BBRSC, MRC, and the Wellcome Trust, possibly supplemented by others (e.g. the US NIA).

A similar joint initiative would be required to enable exploration of these UKHLS and related other prospective study resources for the social sciences in an integrated way. Possibilities might include: establishing a major research centre (perhaps following the very successful MRC model of the Social, Genetic and Developmental Psychiatry Centre at the Institute of Psychiatry); establishing several ‘clusters’ of researchers, led by Professorial Fellows and including generous post-doctoral funding; substantial investment in training to build a cadre of interdisciplinary researchers; an open call for cross-Council funding, drawing on the experiences of the joint ESRC/MRC Society, Social Behaviour and the Neurosciences initiative; using an improved ‘speed dating’ approach to broker collaborations across the spectrum of relevant disciplines, drawing on the experiences of the cross-Council ageing initiative; and sponsoring a high profile interdisciplinary seminar series. The funding for these developments (to cover more than the UKHLS alone) should, at a minimum, match the current ESRC investment (£17 million over the next five years) in Genomics Centres, which are more concerned with the social science of science than progressing the scientific understanding of social, economic, psychological and health.

It is much harder to set a priority list for the stable characteristics: the list of possible measures is long and requires careful consideration. However, using the criterion of broad importance for many fields of study, I would place high priority on measuring cognitive performance and the five-factor personality traits and then place priority on early choice of at least one measure from each of the differing subject domains, especially if of wider relevance, and perhaps also on introducing varied measurement approaches (but only if truly appropriate).

The ESRC has shown great foresight in obtaining (partial) funding for the UKHLS. The study must be a success and more funding is likely to be forthcoming if the links to health and biology are made strongly. These links will significantly strengthen the scientific value of this study and play a real part in establishing it as a world-class innovative study. The UKHLS should have a key distinctive role of contributing to the social, economic, demographic, and psychological sciences that are in the purview of the ESRC. Further, integration with the health and biological sciences has to be among the major aims of the study. Maintaining an appropriate balance will serve to add distinction to the study and provide major opportunities for bridging and exploring the interplays among these elements. Advances in understanding the interplays of nature with nurture, elaborating key pathways, and exploring the interrelations among the social, demographic, economic, psychological and health are of key scientific and policy importance. The proposals put forward here are intended to ensure that we gain from the inclusion of biomarkers and retain a broad social scientific focus, whilst also providing real benefits for the health and biological sciences.
References


Plomin, R., J.K.J. Kennedy and I.W. Craig. 2006. The quest for quantitative trait loci associated with intelligence. Intelligence ??:??.


Wellcome Trust Case Control Consortium. 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447: 661-678.
Annex – List of people consulted

Many people were generous of their time, ideas and unpublished work in helping to shape my thinking for this report. None are responsible for the content but all deserve thanks.

Extended discussions with:
Chris Bacharach, Jeff Evans and Mike Spittel, NICHD 
Anne Pebley (UCLA) and Narayan Sastry (Rand) – PIs for the Los Angeles Families and Neighborhood Study
Robert and Tess Hauser (Wisconsin) – PIs for the Wisconsin Longitudinal Study
Kathy Mullan Harris and Guang Guo (UNC Chapel Hill) – PIs for AddHealth Study
David Weir (Michigan) – PI of Health and Retirement Study
Jim Stevenson (Southampton) – lead member of the ESRC SRB on biological and social determinants of individual behaviour

Continued engagement with:
Sara McLanahan, Chris Paxson (Princeton), Irwin Garfinkel and Jeanne Brooks-Gunn (Columbia) – PIs for the Fragile Families Study
Dan Notterman (UMDNJ Robert Wood Johnson Medical School & Princeton)
Burt Singer (Princeton)

Helpful conversations, interactions or readings from:
Noreen Goldman (Princeton - PI) and Chris Seplaki (Johns Hopkins) on the SEBAS Taiwan study;
Elena Grigorenko (Yale), Robert Kahn (CCHMC, Cincinnati), Jim Swanson (UC Irvine), Jenae Niederhiser (Univ. of Maryland) and Tim Newman (Cape Town) as fellow members of the Fragile Families DNA Advisory Group;
Tom McDade (Northwestern), Stacy Lindau (Northwestern), Doug Ewbank (U of Pennsylvania);
Richard Suzman (NIA), Max Weinstein (Georgetown), Jim Vaupel (MPI, Rostock), John Cacioppo (Chicago) and other participants at the NAS Workshop on Collecting and Utilizing Biological and Genetic Information in Social Science Surveys;

Useful building blocks for UKHLS:
Meena Kumari, Mike Wadsworth, Margaret Blake, John Bynner and Gert Wagner
Dieter Wolke & Scott Weich – excerpts from bid documents