Policy for use and oversight of samples and data arising from the Biomedical Resource of the 1958 Birth Cohort (National Child Development Study)

Vsn 4 21st December 2014

Policy for use and oversight of samples and data arising from the Biomedical Resource of the 1958 Birth Cohort (National Child Development Study)

1 Background and history – an abstract from Power and Elliot[1]

The 1958 birth cohort (1958BC) or National Child Development Study (NCDS) started out as a study of Perinatal Mortality that recruited just over 17,000 births in a single week (3rd- 9th March) in 1958. It aimed to identify social and obstetric factors linked to stillbirth and neonatal death and its findings contributed to the improvement of maternity services in Britain and to a subsequent reduction in perinatal mortality. Table 1 outlines the historical development of the cohort. Although the original survey was not planned as a longitudinal study, the National Children's Bureau was subsequently commissioned by the Central Advisory Council for Education (the Plowden Committee) to retrace the cohort at age 7 and monitor their educational, physical, and social development. Further surveys took place when children were aged 11 and 16. Followed into adult life, the cohort reached a stage of life that is marked by major transitions - from school or full-time further education to employment (although unemployment was very high), and from dependent status in their family of origin to independent status as heads of new households. A survey at age 23 (1981) was designed to trace these transitions, and in so doing it differed from earlier follow-ups in obtaining information directly from the cohort member (instead of from their parents, usually the mother). In 1985, responsibility for the cohort was transferred to the Social Statistics Research Unit (SSRU) at City University, London. SSRU undertook a survey at age 33 (1991) that included a random one in three sample of cohort participants. In 1998, the SSRU moved to the Institute of Education, London, and became the Centre for Longitudinal Studies (CLS). CLS also houses the 1970 birth cohort study (BCS70), and in 1999/2000 an integrated contact of both cohorts was undertaken to facilitate comparisons between these generations.

In 2002-2003, a **Biomedical Survey** of the cohort (aged 44–45 yr) was conducted, with several collaborating partners, under the Medical Research Council's (MRC) 'Health of the Public' initiative. This was ultimately led by Professors David Strachan (St Georges Hospital), Chris Power (Institute for Child Health) and Heather Joshi (CLS). A total of 12,037 subjects were contacted and 9,377 were successfully interviewed. The primary objective was to obtain critical biomedical information via questionnaires, physical measures and biospecimen collection (blood, urine and saliva) and to use this information to examine how developmental, lifestyle, and environmental factors act throughout the lifespan to influence current ill health, and physiological and psychological function in early middle age. Additional funding from the Wellcome Trust (WT) enabled the creation of a comprehensive DNA repository which included transformed lymphocyte lines on a total of 7,526 subjects, providing a permanent source of DNA on those cohort members.

It has been a central tenet of the 1958BC, throughout its history, that outputs from the project should be used in a wide variety of different ways to further research in the social, educational and health sciences. Cohort members therefore *expect* that their data and samples will be used for the benefit of society in general. The central philosophy underlying the creation of access policies and the design of access structures therefore emphasises maximisation of scientific gain for society within a framework that guarantees the security and privacy of individual cohort members.

Table 1: 1958BC, dates of contact, sample size, and funding sources (adapted from Power and Elliott, 2006)1

Survey	Year[a]	Age (yr)	Data from	Sample targeted[b]	Sample achieved[c]	Funders
PMS[d]	1958	0[e]	Mother and medical records	17 634	17 416	National Birthday Trust Fund
Sweep 1	1965	7	Parents; school; tests; medical exam; cohort member	16 727	15 425	Department of Education and Science (DES)
Sweep 2	1969	11	As sweep 1	16 754	15 337	Social Science Research Council
Sweep 3	1974	16	As sweep 1 + census	16 901	14 647	DES; Department of Health and Social Security (DHSS)
Exams	1978	20	Schools attended at age 16 yr	14 647	14 331	DES
Sweep 4	1981	23	Cohort member; census	16 482	12 537	DHSS, DES; Department of Employment; Manpower Services Commission; Department of the Environment (DOE)
Sweep 5	1991	33	Cohort member; spouse/partner; children[f]; children's mother[f]	16 240	11 407	Economic and Social Research Council (ESRC); Department of Health (DH); Department of Social Security (DSS); Employment Department; DES; DOE; Transport and Road Research Laboratory; Health and Safety Executive; US National Institute of Child Health and Development
Sweep 6	2000	42	Cohort member	16 240	11 419	ESRC; Office of National Statistics; Department of Education and Employment; DSS; DH, Scottish Executive, Basic Skills Agency); International Centre for Child Studies
Biomedical Survey[g]	**2003**	**45**	**Cohort member**	**16 078**	**9 426**	**Medical Research Council; Wellcome Trust**
Sweep 7	2004	46	Cohort member	16 012	9 531	ESRC; Department Education and Skills
Sweep 8	2008	50	Cohort member	16 014	9 790	ESRC
Sweep 9	2013	55	Cohort member	11 500 (Estimate)	9 200 (Estimate)	ESRC

[a] Fieldwork often extended over more than 1 yr.

[b] All those born in Great Britain during the week 3-9 March 1958, still living in Britain at that sweep (+ those included from outside Britain during the childhood surveys).

[c] All those in the target sample that actually participate (at least one survey instrument partially completed).

[d] PMS = Perinatal Mortality Survey.

[e] At birth

[f] For a random sample of 1 in 3 cohort members, information was collected directly from 3,008 children of cohort members, 2,588 mothers (could be the cohort member, their spouse, or their partner) and a further 1,270 children that were children of the spouse/partner but not of the cohort member.

[g] The Biomedical Survey was a collaboration between C Power (Institute of Child Health, London); D Strachan (St George's Hospital Medical School); Centre for Longitudinal Studies, National Centre for Social Research.

2 Data and samples

2.1 What and where?

The full 1958BC resource may conveniently be viewed as consisting of four principal elements

1) Data collected for studies focussing on perinatal mortality and on educational, physical and socio-economic development collected in surveys from "PMS" to "sweep 6" and "sweep 7" to "sweep 9" (this will be referred to as the **Social Studies resource**);
2) Information/data pertaining to measures of health and disease in the Biomedical Survey (the **Biodata resource**);
3) Blood samples and transformed cell lines from the Biomedical Survey, the DNA extracted from the blood and cell lines, and the extensive genotypes that have been generated from the DNA (the **Genetic resource**);
4) Biomedical samples (blood, urine and saliva) collected in the Biomedical Survey that are used for purposes other than generating DNA (the **Biospecimen resource**). It is these terms that will be used throughout this document.

The resource represented by the combined ***Biodata***, ***Genetic*** and ***Biospecimen resources*** will collectively be referred to as the **Biomedical resource**.

2.1.1 The Social Studies Resource and Biodata Resource

The *Social Studies resource* and the *Biodata resource* are stored and accessed via UK Data Service (UKDS) at the University of Essex (http://ukdataservice.ac.uk/). Details of the data available in the *Social Studies* and *Biodata* resources can be explored on the UKDS website or using the study Data Dictionary on the CLS website. Active guidance on the availability and use of these data is provided by the Centre for Longitudinal Studies (CLS), under the management of Mr Jon Johnson, overseen by Professor Alissa Goodman (currently Principal Investigator of the 1958 Birth Cohort). Access mechanisms are described in section 3.

2.1.2 The Biospecimen Resource and Genetic Resource (except GWA data)

The active repositories of DNA (native and cell line DNA) and of the cell lines themselves are stored in alarmed freezers and cryostores at the University of Bristol in the ALSPAC laboratories under the management of Dr Sue Ring. Here, the term *active* implies that these repositories are used directly to supply users. Once a DNA award has been made and a Material Transfer Agreement (MTA) signed, DNA plates are prepared by Bristol and shipped to researchers. The cell lines are backed up by an *inactive* archive held at the European Collection of Cell Culture (ECCAC) , Public Health England at Porton Down, Wiltshire. Access mechanisms are described in section 3.

2.1.3 The Genetic Resource (GWA data only)

Most genotypes so far generated on 1958BC participants have been created by a small group of leading research consortia who have carried out genome wide association (GWA) scans: (i) the Wellcome Trust Case Control Consortium (WTCCC); (ii) the Type 1 Diabetes Genetic Consortium (T1DGC); (iii) the GABRIEL Consortium; (iv) the METABOCHIP Consortium; and (v) the IMMUNOCHIP Consortium. Details of the GWA data sets generated from the 1958BC are available from

the European Genome-phenome Archive (EGA) website and are summarised in Table 2. Data from smaller genotyping projects are held in Bristol. Access mechanisms are described in section 3.

Table 2: GWA data sets available from 1958BC participants

Chip Used	STUDY							Total (per chip)
	WTCCC1	WTCCC2	T1DGC	GABRIEL	METABOCHIP	IMMUNOCHIP	HLA	
Affymetrix 500K	**1504**	-	-	-	-	-	-	**1504**
Infinium HumanHap 550K v3	**1436***	-	**2604**	-	-	-	-	**4040**
Illumina 15K Custom chip	**1504**	-	-	-	-	-	-	**1504**
Affymetrix v6	-	**3000**	-	-	-	-	-	**3000**
Illumina 1.2M	-	**3000**	-	-	-	-	-	**3000**
Illumina HumanEXome-12v1_A-GenCall, zCall	-	-	-	-	5841	-	-	**5841**
Illumina Human 610-Quad	-	-	-	8141**	-	-	-	**8141**
Illumina ImmunoBeadShip – Illuminus, GenoSNP	-	-	-	-	-	6812	-	**6812**
Dynal RELI SSO assay	-	-	-	-	-	-	6662	**6662**
Total (per study)	**4444**	**6000**	**2604**	**8141**	**5841**	**6812**	**6662**	**94504** **40504**

*** Not yet deposited with the EGA website April 2014**

****Full number not currently available on the EGA website April 2014**

2.2 Dissemination of information about resource availability

In order to optimise the value of the research infrastructure provided by the 1958BC, the resources are well advertised and are widely known, both nationally, and across the world. Information about the cohort and its associated resources can be found on the websites for: the Centre for Longitudinal Studies (www.cls.ioe.ac.uk/), University College London, Institute of Child Health (http://www.ucl.ac.uk/ich/research-ich/mrc-cech/cohort-studies/1958) MRC (www.mrc.ac.uk), WT (www.wellcome.ac.uk), and UK Data Service (http://ukdataservice.ac.uk/). At present, the website of the University of Leicester (http://www2.le.ac.uk/projects/birthcohort) provides information about accessing data or biosamples collected under the Biomedical Sweep. Detailed information about all data collection sweeps, including data dictionaries, is accessible at: www.cls.ioe.ac.uk/page.aspx?&sitesectionid=117&sitesectiontitle=Surveys+and+documentation.

3 Oversight and management of access to data and samples

The Access Committee for CLS Cohorts (ACCC) has responsibility for the oversight and management of access to data and samples from three British Birth Cohorts run by CLS; the 1958BC, the 1970 Birth Cohort, and the Millennium cohort. It also provides oversight of access to the UK Twins Study. Membership of the ACCC is detailed in Appendix A and in the Terms of Reference (TOR) in Appendix B. The ACCC reports to the Strategic Advisory Board (SAB) of CLS. At the time of writing (March 2014) this new Board is in the process of being constituted – its first meeting will take place in May 2014. The composition of members is listed in Appendix C. One of the first tasks of the new SAB will be to ratify new Terms of Reference for the ACCC – at present, a provisional version, which has been extensively discussed by the ACCC, the funders and CLS, appears in Appendix B and on the ACCC website (http://www2.le.ac.uk/projects/birthcohort/oversight-committee) . This will be replaced by the definitive version, once that has been agreed with the CLS SAB.

Membership of the ACCC (Appendix A) consists of research professionals from a variety of backgrounds including epidemiologists; clinical scientists; social scientists. The expertise of the Committee subsumes genetics, psychology and economics. The ACCC is served by a secretariat and a Technical Review Team (Appendix A) that have detailed knowledge of the data and samples available from the 1958BC: Dr Sue Ring (head of the DNA/cell line resource at the ALSPAC Laboratories in Bristol University); Mr Neil Walker (expert in the large scale genotyping resources at the Diabetes and Inflammation Laboratory in Cambridge who also has extensive knowledge of the GWA data set structures at the EGA); Mr Jon Johnson (the informatician responsible for the 1958BC Social Studies resource and the Biodata resource, based at CLS).

3.1 Access Mechanisms

Although the ACCC has formal responsibility for overseeing access to all data and samples from the 1958BC, the mechanism of review and data release varies depending on the class of data requested (Social Studies, Biodata, Genetic, Biospecimen or Linked) and the administration and release of some of the data is devolved to other bodies. A summary of the oversight and management is detailed below for each type of resource.

3.1.1 Social Studies Resource and Biodata Resource

Data in the *Social Studies Resource* is stored with, and released by, the UK Data Service (UKDS) with direct oversight by Professor Alissa Goodman at CLS. Users wishing to apply for access to the data must register with UKDS (http://ukdataservice.ac.uk). If they are attached to a UK institution of higher or further education (UKHE/FE), they are able to register using the username and password issued by their institution (using the ATHENS authentication system). Other UK users who are not attached to an organisation which is part of the UK Access Managements Federation (UKAMF) - and users who are not based in the UK - can still access the resource but need to apply for a UK Data Archive username and password. Due to data redistribution licence agreements non-UK users have some restrictions on the data accessible, but the granting of Special Licences has recently (February 2014) been extended to all residents of the European Economic Area (EEA).

Registered users of UKDS are able to download data upon registration of a proposed utilisation plan, and formal agreement to the conditions specified in an End User Licence (Appendix D) and where necessary, a Special Licence. (Appendix E) Data which are particularly sensitive or pose a significant disclosure risk – *e.g.* fine level geography - are *only* available under Special Licence. A Special Licence imposes certain restrictions on the handling and usage of the data, and enables identification, oversight and audit of those using the data. It also enables confirmation that potential users are *bona fide* scientific researchers. Versions of the datasets that require such a clearance are marked as 'Special Licence' datasets on the UKDS website.

Access to the *Biodata resource* is also administered through UKDS and is managed in the same way as the *Social Science resource*. At present, data collected from the Biomedical Survey that are held by the UKDS can *only* be obtained under Special Licence, but this is being kept under review. Requests under Special licence which raise particular scientific or ethical concerns are reported to the ACCC for discussion and adjudication. Prof Alissa Goodman at CLS oversees access to 1958BC data via Special Licence and she and Mr Jon Johnson, also at CLS, provide advice on obtaining and completing the required Special Licence. Any requests for 1958BC data from UKDS which raise particular scientific or ethical concerns are referred to the ACCC for discussion or adjudication. This reflects the fact that although the vast bulk of requests for 1958BC data from UKDS are processed entirely by Alissa Goodman and Jon Johnson, the ACCC nevertheless maintains formal oversight of these awards. It is just that with mutual agreement, the ACCC has delegates this specific responsibility back to CLS and their 1958BC staff. The ACCC receives updates on the number of awards that have been made via this route. This mechanism based on delegation is precisely equivalent to the delegation of the overall responsibility for oversight of access to 1958BC data and biosamples to ACCC by CLS and the funders (formally via CLS SAB). Crucially, this implies that if serious problems were perceived to have arisen with either mechanism, the relevant delegation could formally be revoked and alternative arrangements instituted very rapidly.

3.1.2 Biospecimen and Genetic Resource (except GWA data)

Access to the Biospecimen resource and the Genetic resource (except GWA data sets held by the EGA) is administered directly by the ACCC. The ACCC also reviews all applications for EGA genotypes that require linkage to any other 1958BC data set. All applications for access are made on a standard form (Appendix F) which is available for download at www2.le.ac.uk/projects/birthcohort. Working in partnership with the ALSPAC project, an entirely web-based application mechanism is currently being developed, but this is unlikely to be implemented before 2015. So, at present, any research

group (UK, or overseas) wishing to make use of these resources must apply for access by downloading and completing the standard form. The information collected on the application form is aimed at permitting assessment of a series of assessment criteria that are applied to each application:

Assessment criteria

(1) Has the application been submitted by *bona fide* researchers?
(2) Does the application violate (or potentially violate) any of the ethical permissions granted to the study or any of the consent forms signed by the participants or their guardians?
(3) Does the application run the risk of producing information that may allow individual cohort members to be identified?
(4) Does the application run a significant risk of upsetting or alienating cohort members or of reducing their willingness to remain as active participants in 1958BC based research?
(5) Does the application address topics that fall within the acknowledged remit of the 1958 project, as understood by participants?
(6) Does the application request access to an infinite resource (data or cell line DNA) or a finite resource (whole blood extracted DNA, blood, saliva and urine)?
    - If the request is for a finite resource, then the application is seen as being in competition with other potential applicants (both current and future) and the quality of the science is reviewed formally (if necessary using independent external reviewers). Please see our Biosample Strategy Guidelines for full details. (Appendix G)

***The ACCC is co-ordinating a call for proposals to utilise samples from the finite resource during 2014 please see the website for more details (**http://www2.le.ac.uk/projects/birthcohort/1958bc/open-calls**).**

All applications for data and samples submitted to the ACCC are first reviewed by a Technical Review Team (see Appendix A) to ensure that all basic technical and scientific questions are identified and addressed and all problems highlighted, and where possible, resolved, *before* an application is passed on to the ACCC for formal consideration. The ACCC then provides a formal ratification mechanism for those applications that are straightforward and problem free, and a forum for discussion, problem resolution and appropriate rejection/acceptance for those applications that are other than straightforward. The ACCC also has responsibility for reviewing and developing new access policies to reflect rapid changes in the scientific, social and ethico-legal underpinning of this important area of bioscience.
The ACCC meets monthly by teleconference with face-to-face meetings 3 times per year. New applications and amendments are guaranteed to receive technical review and potential consideration at the next meeting, provided they are submitted at least one full working week ahead of the date of the meeting. However, if the technical review team identifies one or more substantive technical problems, applicants are rapidly informed of those problems (by email and/or by phone) and are warned that ACCC review cannot take place until all problems have been resolved. Under these circumstances, an application is only kept on the agenda for the *next* ACCC meeting if the applicants are confident that they can resolve the issues and submit a modified application *ahead* of that meeting.

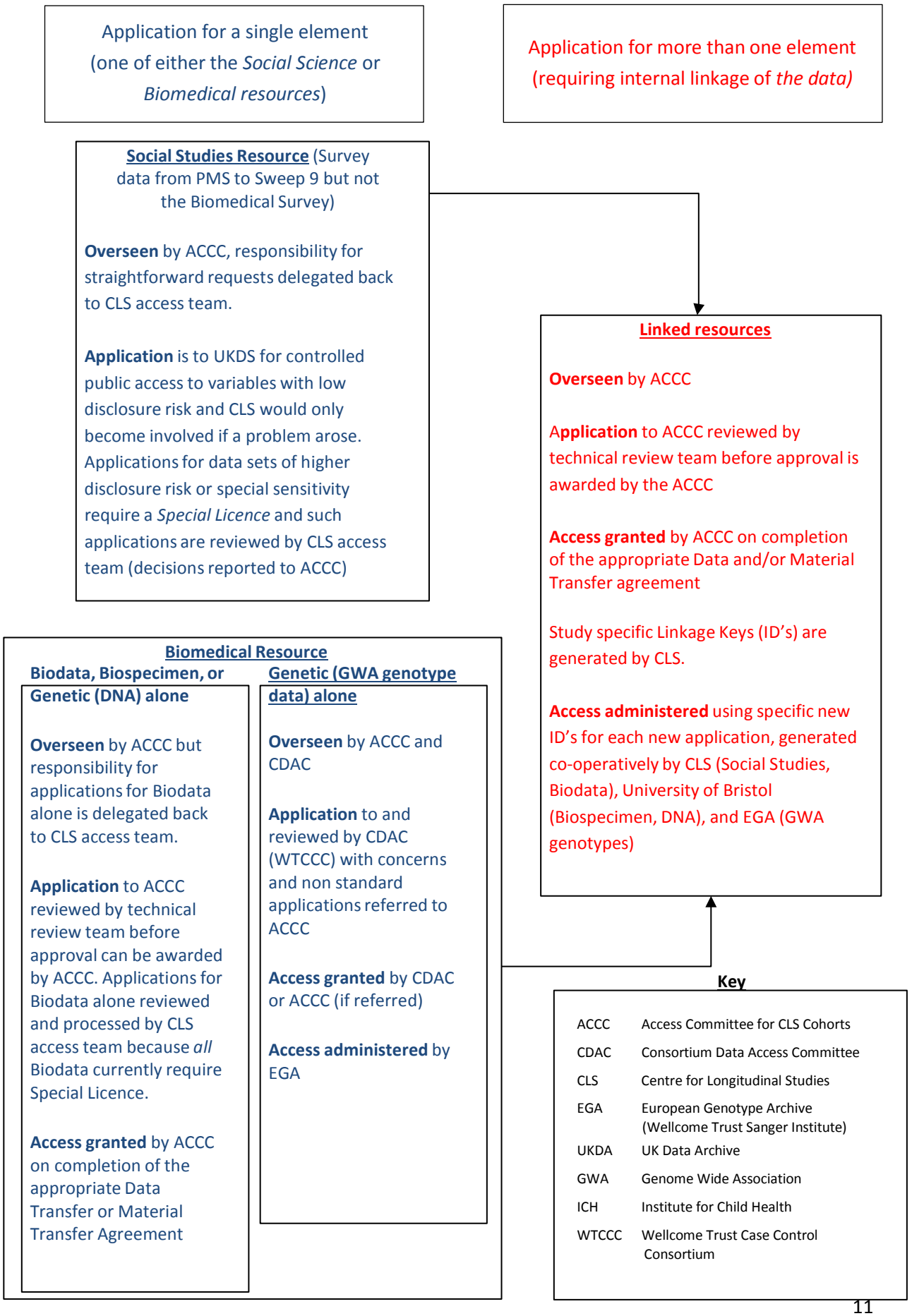
3.1.3 The Genetic Resource (GWA data only)

In the same way that control of access for straightforward applications to the Social Science and Biodata resources held by UKDS are delegated back to Alissa Goodman at CLS, access to the GWA data sets archived on and made available by the European Genome Phenome Archive (EGA) in Cambridge (at EBI on the Sanger Campus) is devolved to the Consortium Data Access Committee (CDAC) of the Wellcome Trust Case Control Consortium (WTCCC). The CDAC is formally convened by Wellcome Trust, under a secretariat led by Dr Audrey Duncanson. (Appendix H) These genotypes are always released with two items of supplementary data: (1) sex; (2) region of residence (a coarse categorisation based on dividing Great Britain into 12 large regions). These additional items of data are essential to permit effective genetic analysis and are agreed not to risk identification of individual participants in the 1958BC. Access to all other genotypes (mainly small sets of genotypes created a variety of *ad hoc* users) is administered directly through the ACCC as detailed above.

Because the WTCCC CDAC deals only with anonymised genotypes, judgement is based solely on Assessment Criteria 1-5 (see above). If applicants to CDAC for access to the 1958BC resource request *anything* beyond the anonymised genotypes from the GWA scans detailed in Table 2, they are necessarily referred on to the ACCC directly. Similarly, if concerns are raised under criteria 1-5 but the CDAC does not feel that the concerns are severe enough to warrant outright rejection, the application will be discussed with the Chair of ACCC and if necessary referred on to that committee for formal review.

Applications for biodata alone, DNA (or gentotypes) alone or biospecimens alone will be reviewed as described above (see Appendix G) for our 2013 guidelines for use of and access to biosamples). If and when approval is awarded, a Data Transfer Agreement (DTA) or Material Transfer Agreement (MTA) (Appendix I), or both, will be signed. With the exception of GWA data released under CDAC, the DTA for all data in the *Biomedical resource* is the Special licence. Where applicants require access ***solely*** to the GWA data overseen by EGA, the applicant signs off on the conditions implied by the acceptance documentation required by WTCCC CDAC (Appendix J). All of these documents indicating access to data or biosamples include a non-negotiable requirement to return new scientific data (*e.g.* genotypes, or new variables) to a central 1958BC scientific archive to be held at Bristol, which will increase the value of the 1958BC resource to the scientific community as a whole. Under all but the most unusual of circumstances if a potential research user works in a legal jurisdiction in which the wording of the MTA or DTA or the requirement to return newly created data is problematic, that is their problem to solve. The fundamental principles are immutable and users that cannot satisfy them because of local jurisdictional problems (legal or ethico-legal) will be unable to access the 1958BC resource.

Figure 1: Oversight and management of access to data and samples.



3.1.4 Linked Resources

When users wish to access more than one type of data, this can potentially increase the disclosure risk, and so such applications demand careful linkage of the relevant data that are required so as to enable secure analysis at an individual level. In the extreme case, a research group may require access to Social Science data, phenotypes from the Biomedical resource, GWA genotypes, cell-line DNA, and blood samples. All of these must be linked together at an individual level in a manner that prevents the data/sample providers (CLS, University of Bristol, EGA) from having access to everything at once because this would violate the governance principles originally established for the study. In addition, it means that end-users (research applicants) are unable to identify individual participants, either from the resources they have been awarded, or by joining their data together with another end-user who has been awarded a different set of data.

In order to achieve this underpinning level of security, the data are prepared and released from a modular platform involving separation of the various study resources. Thus, the data are split into three domains: (1) *Social Studies resource*/*Biodata resource*; (2) *Genetic resource* (other than GWA genotypes from EGA) / *Biospecimen resource*; (3) GWA genotypes from GWA

The first domain is under the management of CLS, the second under the University of Bristol and the third under the EGA. The ACCC sanctions these three centres to exchange information to enable linkage (*i.e.* transfer of identifiers) but without exchange of *actual scientific data*. The basis of data management relating to the *Biomedical resource* (deriving from its originating governance principles) is that large subsets of the data bases fundamental to domain 1 and to domains 2/3 must *never* be brought together in a way that enables the two to be linked directly. CLS is responsible for the governance of study IDs and ensures that this fundamental principle is not violated. Crucially, CLS cannot link to data that are in domains 2 and 3 because CLS does not have physical access to the data in Bristol and will not apply for access to the genotype data at EGA. Furthermore, neither Bristol nor EGA can link to data that are in CLS's domain because they do not hold the relevant linkage keys (they *could* theoretically download non-Special Licence data from ESDS, but they wouldn't know which row of data related to which row of data in their own databases). Table 3 and Table 4 illustrate which data sets CLS and Bristol could individually link together, without the assistance of the other centre: crucially, neither centre can link domain 1 to domains 2 or 3 without cooperation.

Once an application has been ratified that requires linkage between two or more domains, a new study-specific ID will be created which can be mapped to the various generic IDs in each of the three domains. The required domain-specific data will be released to the end-user indexed by this study-specific ID. End-users, themselves, will then link the data together across domains. Generic IDs will never be released to end-users, and data sets will be re-ordered when study-specific IDs are attached. The generation of re-indexed and re-ordered data sets will be jointly managed by CLS, Bristol and EGA.

Table 3: Data resources CLS are able to access together, in a linked form, without cooperation from Bristol

	Social Studies resource (data from UKDS)	**Biodata resource** (data from UKDS)	**Genetic resource** (GWA genotype data*)	**Genetic resource** (extracted DNA, cell lines and non-GWA genotypes**)	**Biospecimen resource** (not extracted DNA or cell lines)
Social Studies resource (data from UKDS)	Same resource	**YES**	NO	NO	NO
Biodata resource (data from UKDS)		Same resource	NO	NO	NO
Genetic resource (GWA genotype data*)			Same resource	NO	NO
Genetic resource extracted DNA, cell lines and non-GWA genotypes**)				Same resource	NO
Biospecimen resource (not extracted DNA or cell lines)					Same resource

*Held at EGA **Held at University of Bristol

Table 4: Data University of Bristol are able to access directly, in a linked form, without cooperation from CLS

	Social Studies resource (data from UKDS)	**Biodata resource** (data from UKDS)	**Genetic resource** (GWA genotype data*)	**Genetic resource** (extracted DNA, cell lines and non-GWA genotypes**)	**Biospecimen resource** (not extracted DNA or cell lines)
Social Studies resource (data from UKDS)	Same resource	NO	NO	NO	NO
Biodata resource (data from UKDS)		Same resource	NO	NO	NO
Genetic resource (GWA genotype data*)			Same resource	**YES**	**YES**
Genetic resource extracted DNA, cell lines and non-GWA genotypes**)				Same resource	**YES**
Biospecimen resource (not extracted DNA or cell lines)					Same resource

*Held at EGA **Held at University of Bristol

3.1.5 Standard Classes of Application

The various data/biosample resources and the application processes outlined above, define a series of 'standard classes' of application that - unless they are associated with specific ethico-legal concerns arising from particular issues other than the types of data they are requesting to bring together – do not require additional ethics clearance. That is, uncomplicated applications falling into any of these defined standard classes will fall under the overall ethics clearance awarded to the 1958BC Resource as a Research Tissue Bank (RTB).

These classes reflect applications for any combination of the five classes of data/biosamples listed in tables 3 and 4 – including anything between a single class on its own, right up to all five classes combined. Crucially, applications involving linkage of two or more classes only fall under the Research Tissue Bank favourable opinion if the data/biosamples are linked and released in the manners described in this document.

To date, no formal record has been made of the particular combination of resources that are requested by each application and the class into which their application actually falls. Settings in which particular ethico-legal issues mean that a given application requires additional ethico-legal clearance beyond that provided by the Research Tissue Bank favourable opinion should, in the first place, be identified by applicants. However, the ACCC may decide that an application that has not been identified as requiring additional ethico-legal clearance *does* in fact require such clearance. If there is then an unresolvable disagreement, the question will be referred up to the CLS Strategic Advisory Board and their decision will be final.

3.2 Finite versus infinite resources

It is critical that over-zealous assessment of the science is not seen as providing an excuse to hinder access to groups that may wish to develop research programs that are strong and worthwhile. In particular, it is important that there is no perception that applications might be rejected because they appear to be in conflict with, or compete with, those being run by the principal investigators or the scientists involved in the oversight of the 1958BC. Consequently, provided access is sought only to resources that are effectively infinite in extent (*e.g.* data or cell line DNA), the assessment of applications is based solely on: (1) the scientific *bona fides* of the applicants; (2) an evaluation ensuring that the proposal violates no ethico-legal principles of the 1958BC (as laid out in consent forms, information sheets, the wording of ethical clearances *etc*); and (3) a judgement that there is no risk that the proposal might "harm" individuals in the cohort, or the cohort as a whole. These pivotal issues are addressed by criteria 1-5 in section 3.1.2 (above).

On the other hand, if the request is for a *finite* resource (*e.g.* biospecimens other than cell line DNA) assessment criterion 6 is triggered, demanding that a formal scientific review has been, or is to be, undertaken. Circumstances differ so markedly between individual applications that it is impractical to devise fully codified rules that will cover all eventualities. It is therefore important to allow the ACCC to make decisions about the level of assessment that is required in each case. In order to make the recommended process as transparent as possible, the ACCC commissioned a formal review of the strategy that might be adopted to deal with applications for finite biosamples. The Biosample

Strategy Guidelines are attached in full to this document as Appendix G. Its summary section is as follows:

SUMMARY

The 1958 tissue samples are a valuable resource but there are limitations regarding their suitability for some assays due to the sample processing history. Recommendations for ACCC for approving use of the samples are:

- *Scientific strength of the proposal must justify use of 1958 cohort samples.*
- *Evidence must be provided to show methodology is appropriate given the processing history of the samples. Eg. Evidence from published literature or pilot data generated on samples processed in a similar manner.*
- *The assay test platform should have proven quality assurance measures in place.*
- *The methodology should include measures to ensure the quality of any remaining sample is not jeopardised and can be used in further assays which can be used on freeze thawed samples.*
- *At least one aliquot of each sample type should be reserved for future global discovery projects.*

In addition, all applications for a finite resource will require formal peer review of the proposed science. This is because all such applications are effectively in competition with other applications – present or future – for the same samples. *As a minimum*, if the request is for a finite resource and the application is not funded by a grant that has been subject to formal (and successful) peer review by a competent authority (*e.g.* MRC, Wellcome Trust, other major health charities) the ACCC will send the application out for external peer review by *at least two* independent assessors.

Furthermore, the ACCC reserves the right to send *any* application out for independent peer review regardless its previous review history. But, if the ACCC eventually rejects an application that has previously been subject to successful peer review (and a grant awarded), and that rejection is based *solely* on its scientific rigour, the decision will be referred to the CLS Scientific Advisory Board for ratification. It is important that the ACCC is not seen to be denying access to resources on the basis of scientific judgements that may be questioned.

In association with the funders and CLS, the ACCC is currently advertising a six month call for applications for any use of the 1958BC biosamples. No formal judgements on any of the applications received during this period will be made until the call has closed. This will enable us to gain an effective overview of the range of applications that might exist and to combine together requests that require similar samples so we avoid thawing and dividing aliquots that could be used for several purposes simultaneously. It is anticipated that such calls will likely become a regular feature of managing access to the 1958BC biosamples.

All data are viewed as an infinite resource. In discussion with the funders it was formally agreed that decisions to award or not award an infinite resource would be judged entirely on their own merits and would take *no account* of other applications that may be similar. However, if similar applications do appear the ACCC may attempt to encourage collaboration between the groups, unless that is seen as jeopardising confidentiality.

3.3 Applications from researchers associated with the biobank.

Applications for access to data or samples for research purposes from scientists or other professionals directly associated with the 1958BC will be dealt with in precisely the same way as applications from anybody else. This will include all scientific investigators, technical staff and academic staff at Bristol University, CLS or EGA, as well as any other staff involved in the oversight of data samples and access that may happen to be based elsewhere.

Any individual on the ACCC who is actually involved in an application for an infinite resource may contribute information to the discussion about that application but they are excluded from the decision making process. This operational policy has been adopted because the ACCC is small and the number of applicants on some applications is so large, that the quality of discussion about some access requests would be seriously impaired if applicants were prevented from contributing to the discussion underpinning a given decision. It is the applicants concerned who will typically know the most about the area concerned. For applications involving finite resources, however, applicants will be entirely excluded from discussion and decision making. It is for this reason that decisions regarding a potentially large number of simultaneous applications for finite resources – for example under the 2014 call for biosample applications – will be reviewed by a panel of experts that will be independent of ACCC (though it may include some ACCC members) and will be chosen to be as non-conflicted as possible for the competing applications.

Should the process of producing linked data and/or samples to meet a specific access request require that any one centre simultaneously holds what is seen to be an unreasonably large amount of *mutually linked* data and/or samples (see section 3.1.4), alternative linkage mechanisms will be explored and adopted (if necessary using other third party institutions to undertake the linkage). To date, this has never been required. Mutual linkage refers to a situation in which a block of data and/or samples (spanning more than one element of the 1958BC resource [see section 3.1.4]) are all linked together on the same identifier or on a series of identifiers that can themselves be linked. Should concern arise that excessive mutual linkage will occur if a particular application is awarded it is a responsibility of the ACCC to ensure that all reasonable attempts are made to find an effective solution. If no acceptable resolution can be determined, the proposal concerned will necessarily be modified or declined. But, if the ACCC believes that this rule needs to be triggered, it will refer the decision to the CLS SAB for ratification: the security of the cohort is paramount, but it is important that individual members of the ACCC do not attempt to use this rule to hinder or block reasonable applications.

3.4 Conflicts of interest and appeals

Any applicant who wishes to appeal the decision of the ACCC or to appeal their designation as being conflicted can apply to the CLS SAB, but this will require a documented (self-contained) description of all of the relevant background and a formal justification for why the decision that has been taken is being appealed. To date, no such appeal has ever been required.

3.5 Security

In the real world, regardless of all of the security features that may be built into the data access mechanism, someone with “malevolent” intent could potentially link up data and samples obtained

under the formal oversight structure with data that have been in the public domain for many years via ESDS. This could potentially allow them to accrue enough information to identify some individual cohort members. One of the distinct advantages of restricting access to the *Biodata resource* to Special Licence holders is that we can ensure: (1) those scientists who hold 1958BC biomedical data or genotypes are explicitly informed* that they are not permitted to try to identify individual participants; (2) these same scientists are compelled to *sign off** on a statement that says they will *not* try to do this; (3) if any attempt is made to do this (successful or otherwise), appropriate *sanctions* can be applied that will seriously impede the capacity of offenders to undertake future research. At present, such sanctions are necessarily limited to prohibition of future access to data/samples under* ** ACCC, CDAC or UKDS. But, for some time, the broader bioscience community has been discussing global electronic key-based systems to indicate the scientific *bona fides* and *on-going good standing* of international scientists who wish to use any major bioscience resources that may be registered under the putative system. This will offer the potential to deny access to scientists who seriously misuse data or samples - in any way whatsoever – to a range of key databases across the world.

*Terms and Conditions for potential collaborators utilising research materials from the 2002-2004 Biomedical Survey of the 1958 British Birth Cohort, (Version 6a, 5 January 2005, www.b58cgene.sgul.ac.uk) (8) states: *Applicants receiving data are required to agree that they will not attempt to identify individuals within the cohort and they will need to sign the Centre for Longitiudinal Studies Code of Practice to formalise this agreement. Applicants receiving biological materials will need to sign a Materials Transfer Agreement which incorporates the principles of the CLS Code of Practice (see Appendix I).*

**British Birth Cohort (1958BC) Material Transfer Agreement for DNA or Biospecimens (version 6) 12 Feb 09 Terms and conditions.(9) states: . *Not attempt to trace, contact or identify any individual member of the 1958 Birth Cohort or to recruit any cohort member to take part in any other survey. (Appendix I)*

**CDAC Data Access Agreement, Terms & Conditions (3) states: *Agree to preserve, at all times, the confidentiality of information and Data pertaining to Data Subjects. In particular, undertake not to use, or attempt to use the Data to compromise or otherwise infringe the confidentiality of information on Data Subjects and their right to privacy.*

4 Uses and users of the 1958BC resources

4.1 Legitimate uses and users

The 1958BC resources are available to *bona fide* research scientists (biomedical, psychosocial, educational or ethico-legal) from anywhere in the world that have successfully applied to either the ACCC, CDAC oversight committee (see above) or have been permitted access via UKDS/CLS. The resource is used widely for research in genetic and genomic epidemiology – in particular as a platform for genetic association studies. To that end, it provides a source of subjects that have been well characterised (phenotyped) for a wide variety of quantitative complex traits and have also been carefully assessed for a wide variety of sociodemographic, socioeconomic and life style variables that are outcomes in their own right, and may be important determinants in relevant causal pathways leading to disease. It also provides a source of “cases” and “controls” for common binary traits. Finally – and possibly most importantly – it provides a geographically representative sample of British people (of primarily European origin) that represents the premier source of national controls that can be used in a wide variety of genetic case-control studies. For example, the 1958BC was used as one of two sets of national controls in the Wellcome Trust Case Control Consortium, that studied more than 12 complex diseases and was voted international “Research Leader of the Year, 2007” in the Scientific American SciAm 50 awards. On the basis of this contribution alone, the 1958BC Biomedical Survey has already played a major role in national and international bioscience.

4.2 Restrictions on uses and users

There are a small number of restrictions on the uses to which the 1958BC resource may be put.

(1) Data can only be obtained in anonymised form, and must not be used, reported or published, in any way that could potentially lead to the **identification of individual cohort members**.

(2) The consent forms (see appendix K) restrict use to **non-commercial purposes**. For example, in relation to blood samples obtained in the Biomedical Survey, they contain the following statement: *"I understand that the blood samples and related information will be coded and used anonymously for non-commercial research purposes only, and will not be tested for HIV."*

Unfortunately, the term *"non-commercial research purposes"* is ambiguous, and at present we do not really know how individual cohort members interpret it. Consequently, the Access Committee has necessarily adopted a conservative working definition that says that we will allow no use that is led by a commercial organisation, could lead on to a commercial product or benefit, or could lead to a commercial organisation gaining control or ownership over some part, or derivative, of the 1958BC resource. However, we *do* allow researchers to use commercial biotechnology firms to undertake their genotyping (or other high-throughput bioassays) because this is the standard approach used across bioscience (most individual academic institutions cannot afford the infrastructure), and full control and ownership of samples and data remain with the researchers. We also allow *bona-fide* scientists who work for a commercial organisation to use data and samples for purposes that are clearly not aimed at direct commercial gain but rather at providing an infrastructure for new research opportunities in bioscience in the future. For example, a scientist working for a commercial organisation in India has been awarded access to 1958BC data in order to help him to develop the statistical methods and techniques required to develop a "population ancestry map" for India that will be made freely available.

Because of the conservative and relative restrictive position that has been adopted, we have had to turn down some applications that we would have liked to award and that we suspect, if questioned, individual cohort members would willingly have supported. For example, in 2007, Affymetrix offered to pay to fully fund the application of a new Genome Wide Chip across the 1,500 subjects in the WTCCC1 control sample (see Box 1), and to return all genotypes for international access to 1958BC. This would have produced an extremely valuable resource for bioscience and would undoubtedly have enhanced the value of the 1958BC. Furthermore, it would have produced no direct commercial gain, nor ownership/control of any biological resources or intellectual property. But, if the genotyping had been successful, the company would inevitably have wished to use this success in a world leading study in its marketing material, and so this was seen as a *potential* commercial gain.

We suspect that if cohort members were asked the direct question *"would you be prepared for your data/samples to be used by a pharmaceutical (or biotechnology) company to create a new life-saving or life-enhancing medication (or a new way to study genes) even if you were to receive no personal financial return from this?"* most would say yes. That being the case, it would have been preferable if the original consent had been written somewhat differently.

Partly in order to resolve these ambiguities, we are currently considering how best to consent the 1958BC should there be a second Biomedical Survey. But the issues are complex and it is currently unclear what strategy is likely to be adopted.

Restatement of the formal position of the ACCC as of 9th December 2014

Following consideration by both the ACCC and the CLS Strategic Advisory Board, and in light of discussions associated with the successful recent application to the North West Research Ethics Committee for renewal of ethical clearance of the Biomedical Resource of the 1958BC as a Tissue Biobank, the key principles dictating response to commercial applications can now been stated unambiguously. Specifically, research from a commercial organisation will be viewed as acceptable in principle (assuming all other acceptability criteria are met) provided:

1. The research for which access is requested will not, in itself, result in direct commercial gain
2. The research program for which the access is being requested is led by an academic group from a non-commercial institution, with the commercial agency acting as a research partner
3. The commercial agency will gain no ownership – or attempt to claim ownership – of any data or biosamples from 1958BC, or of data or biosamples generated under the permitted research, nor will they place any restriction on the use of any of these data or biosamples by other users

(3) There are also repeated references in the consent forms for the Biomedical Survey (see Appendix K) to the scientific nature of the research to be undertaken. For example, *"I give my consent to storage of frozen portions of my blood sample for use in future medical research studies of the causes, diagnosis, treatment or outcome of disease."*

This is again ambiguous – for example, a small number of commentators have suggested that this should exclude social and educational research while the majority do not. Up to now, the access committee has interpreted this broadly and the ACCC will continue to take this position. Namely, all *bona fide* scientific research will qualify provided it can be demonstrated that the field of study contributes knowledge that may help us better understand health and disease. This immediately encompasses much work of a socio- economic or socio-demographic nature, as well as psychological and educational research. This is because there are so many health-related conditions that have their primary origins in psychosocial, economic or educational determinants or have an impact on one or more of these dimensions. It is perhaps relevant to note that the original vision of the Plowden Committee (see page 1) in converting the NCDS into a cohort study was to investigate educational, physical, and social development.

(4) The patient information sheet for the Biomedical Survey guaranteed that data would not be available for: (i) life insurance purposes; (ii) mortgage applications; (iii) police records; or (iv) HIV/Aids testing. No conceivable *bona fide* research-based application could legitimately involve the first three of these, and so these particular uses should not arise. Any application involving HIV/Aids testing will be declined by ACCC (or CDAC). It is the considered view of the access committee that regardless what may have been written in the information leaflet, if the police or other *bona fide* national authorities were to present a court order demanding information, that request would have to be met. But, in keeping with the spirit of the information leaflet no data or samples will be released to the police or any other branch of

the law enforcement agencies, without a formal court order. This is consistent with advice on confidentiality of medical data from the GMC[2] and is the same policy that has been adopted by UK Biobank.

(5) There are also three specific areas of research that are prohibited because of information given to study nurses during their training for the collection of data and samples in the Biomedical survey. To be specific, they were told that the data, measurements and samples that were to be collected in the survey would *not* be used for the direct study of the genetics of *intelligence, sexual orientation* or *criminality*. Even though participants were not formally given this same information, it was the view of the 1958BC access committee that this prohibition should be respected. In consequence, applications seeking to study these domains have been declined. However, in relation to 'intelligence' as a trait, the Committee has noted that tests of cognitive function involve multiple dimensions that are closely related to intelligence. Following formal discussion in the Committee, it was decided that studies involving measures of cognitive function *do* represent an acceptable use of the 1958BC provided they are carried out to better understand medically important traits such as autism, ADHD and dementia. On the other hand, studies specifically setting out to study 'intelligence' or 'IQ' as primary outcomes are not. Professor Alissa Goodman at CLS is currently drawing up a list of 'concerning' variables that may take an application into one or more of these areas of concern – and none of these variables will be available in linkage with genetic data, DNA or cell-lines without a very strong and convincing justification being put forward.

(6) At the time of the last application for ethics approval, the section about reproductive research read as follows:

The 1958BC resource has been used for some types of reproductive research (e.g. studies of perinatal mortality), and such uses may arise again in the future. If a new proposal in the field was to emerge that was particularly sensitive, it would be considered very carefully by the ACCC. Any application that might be viewed as "sensitive" or "controversial" would trigger concern to the ACCC under one or more of the assessment criteria 1-5 (see above). Given the nature of the resource, and the current state of bioscience, no applications are anticipated in the fields of therapeutic cloning, stem cells or use in animal models.

However, many cohorts are now considering generating induced pluripotent stem cells from blood-derived cell lines. Although no definitive plans to undertake such work on 1958BC currently exist, they may well arise later. If they do, we will approach the Ethics Committee with an amendment specifically addressing this issue. Crucially, no application for such work will be acceptable to the ACCC unless it passes *all* assessment criteria (section 3.1.2). Furthermore, no approval will be given to proposals involving the production of gametes.

4.3 Sensitivities relating to genetic uses of the resource

Among genetic uses of the 1958BC resource, it is used primarily for genetic association studies which are based entirely on anonymised genotypes (single nucleotide polymorphisms [SNPs] or copy number variants [CNVs]). Such studies are rarely associated with intractable ethico-legal problems, unless a disease under study is itself sensitive. The issues that are known potentially to be problematic in such analyses are well recognised and are faced by almost all biobanks – our procedures are consistent with practice in the best biobanks internationally.

Ultimately, any individual may be identified in a genotype data set using relatively few genetic

markers – this (loosely) is the basis of DNA fingerprinting. But given anonymisation of the genotypes this leads to no risk of identification unless either the identity of the person who provided the genotype is already known (in which case identification is irrelevant) or if the genotypes are held with a large number of non-genetic variables that are themselves identifying. For this reason, scrupulous care is taken in the 1958BC to ensure that anonymised genotypes cannot be linked back to data in the *Biomedical* or *Social Studies Resources* (see section 3c).

Considerations about the potential identifiability of genetic research data were made more complex when, in August 2008, Homer et al[4] showed that given an extensive genotype from a given individual, it was possible with high probability to demonstrate whether that subject (or one of his/her relatives) was - or was not - one of a potentially large number of subjects in a group for whom *summarised (group-level) genotypes* were available. This has forensic and non-forensic implications. First, given an extensive genotype from a crime-scene, the police could trawl through summarised genotypes on numerous biobank websites (including 1958BC) until a match was found for their sample. They could then approach that biobank and seek to identify the individual responsible for the match. Second, a family member or a scientist might misuse access to tissue samples or genotypes from an individual in order to determine whether that subject's genotype matched with a summarised genotype provided for a group of cases of a disease which the subject may wish to keep confidential. The potential implications of this paper were discussed in full at an Extraordinary Meeting of the 1958BC Access Committee in October 2008. It was concluded that the potential risks for the 1958BC were very small, but in order to ensure that even this small risk was controlled, summarised genotypes of a nature that could lead to identification would be removed from free access on the web site at St Georges Hospital. In addition, it was decided that, like UK Biobank, if 1958BC was approached by the police - or any other law enforcement data – seeking access to either data or samples, such access would only be provided if required by a Court Order .

Additional concerns for open access data were raised in early 2013 with publication of a paper by Gymrek at al. [1], which pointed out the extent to which inferences based on linking surnames to DNA sequences available on open-access websites (*e.g.* geneology websites) could lead to a risk that *bona fide* applicants for – in our case, 1958BC data - could identify individual participants from their genomic data. But this was not seen as an additional concern for ACCC. Any applicant seeking to identify individuals from genomic data provided by 1958BC will automatically violate the agreement that they have explicitly signed stating they will *not* do this. This will attract punitive professional sanctions and, it has been suggested in several quarters, a risk of criminal prosecutions. If someone apparently *bona fide* is prepared to run this risk for reasons that would appear to produce very little personal gain, there would appear to be very little anybody could do to reliably prevent them from doing *something* unethical – and there is then no real option but to punish them after the event. No action of a nature remotely like the malevolent scenario presented has ever occurred in association with the 1958BC. If it did, the Ethics Committee would be informed.

In summary, it is our considered view that the genetic purposes to which the 1958BC are likely to be put, are unlikely to be particularly sensitive. Where possible, foreseeable risks have been identified and procedures put in place to control those risks. If a genetic analysis is sensitive because of the particular disease that is being studied, then that will trigger one or more of the access assessment criteria and the ACCC will consider the particular application very carefully. Finally, all applicants sign up to a formal professional code of behaviour and, as a final resort, the ACCC will work with CLS/Institute of Education (who hold legal responsibility for 1958BC) and/or University of Bristol who hold legal responsibility for the MTA, to apply strong sanctions – potentially including legal action if warranted – to punish anyone who misuses the 1958BC resource.

4.4 Financial rules for access

In principle, end-users are expected to meet all the costs of DNA handling, specimen transport and data preparation in relation to their study [MTA, Schedule 2 (3)] (Appendix I). They are also liable for the costs of clerical, data processing and/or statistical support incurred in providing research materials [Old terms and conditions for DTA (9)]. But, under current funding, the cost of preparing samples for users is covered by funding from Wellcome Trust and MRC. Users are asked solely to cover courier charges. It is likely that a cost recovery mechanism will be introduced during the next period of funding for 1958BC data access infrastructure (from 1st May 2015).

4.5 Return of data (genotypes, publications *etc*)

Terms and conditions (Material Transfer Agreement Version 6 12 Feb 09) (Appendix I) for potential collaborators utilising research materials from the 2002-2004 biomedical survey now require that:

New data or derived variables

(1) Paragraph 15: It is a condition of access to the samples that all information obtained from the samples (including any derived data, for example on haplotypes) is submitted to University of Bristol for inclusion in the central 1958BC database. The Recipient Institution will keep the ALSPAC laboratory, University of Bristol informed of the Results of the Research. The Recipient Institution will provide the ALSPAC laboratory, acting on behalf of the 1958BC, with a fully documented electronic copy of the Results before publication in any form or within 12 months of the completion of the Research whichever is the sooner. There will be accompanying documentation sufficient to identify the genotype (eg chromosomal location of the genetic variants) or bioassays tested, the interpretation of the coded results, and a brief description of the methods used. The format for this report will be agreed between the Recipient Institution and the ALSPAC laboratory, University of Bristol. Where necessary, the timing of lodgement, and any subsequent embargo on their use by others, will be agreed between the applicants and the ACCC. At the discretion of the ACCC the data may be lodged with the UK Data Archive. Applicants must supply adequate documentation concerning new variables (including statistical programs) to permit their use by others in future analyses of the data.

(2) Paragraph 19: Return or destroy the Material at the end of the project as requested by the ACCC.

(3) Paragraph 8: Communicate promptly and in writing (E-mail is acceptable) to the University of Bristol any information regarding the quality of the Material or problems they may encounter with the Material or errors in the Material.

The ACCC requires that study level results are made available to other users in accordance with contemporary best practice in medical science and taking appropriate account of the ethico-legal restrictions arising from study consent and ethical clearance, and recognizing the potential risks of disclosure of summary level genotypes.[4] If this stipulation is not followed, the ACCC will take this into account in judging future access requests from the responsible applicants.

Reports and publications

(1) Paragraph 18: Provide reports of progress or any other nature as requested by the ACCC, and notify the ACCC of any significant delays in completing the research proposed in schedule 2.

(2) Paragraph 16: The Recipient Institution will acknowledge 1958BC and the funders and will adopt standard publication policies in determining how to reflect (authorship or appropriate acknowledgement) individuals ACCC who have played a substantial scientific role in the

generation of the Material used in a specific publication based on 1958BC data, samples or results. The secretariat to the ACCC must be informed of all research papers based wholly or partly upon the Material.

(3) Paragraph 17: Inform the press offices at the Wellcome Trust and Medical Research Council prior to any media publicity.

If key individuals involved in the construction of the development of the 1958BC Resources and Material are not appropriately reflected in authorship or by acknowledgement, the ACCC reserves the right to take this into account in judging future access requests from the responsible applicants.

Finally, if an application looks as though it may produce results that are controversial or may take use of the 1958BC data/biosample resource close to areas of use that are viewed as unacceptable, the ACCC reserves the right to tell successful applicants that they should submit any papers they produce to Professor Alissa Goodman at CLS (as PI of 1958BC itself) ahead of publication. This is not intended to introduce a significant delay in publication or a threat of publication being prevented but rather to ensure that CLS is then in a position to respond effectively to any queries they may receive from participants, the media or any other bodies or persons.

4.6 Engagement with donors/participants

All cohort members participating in the survey were sent a letter of thanks as soon as possible after the Nurse Interviewer visit. Cohort members also receive an annual birthday card and occasional feedback on key findings and updates on future plans. There is also a participant's website providing information and results of findings at http://www.ncds.info.

4.7 Feedback of clinically relevant results to participants

One of the greatest challenges for those running contemporary cohort studies is how to deal with clinically actionable findings. These largely arise from medical imaging, or from dense genotyping or sequencing of DNA variants associated with a high risks of developing a serious clinical condition that may be remedied by appropriate intervention. To date there has been no medical imaging in 1958BC, but dense genotyping is becoming increasingly common. An important issue faced by 1958BC is that, at the time when the Biomedical Survey was undertaken, international best practice was almost universally to ask participants to sign a consent form agreeing they would never receive back any genotypic information.

But as an increasing number of potentially clinically actionable variants are being detected, most large cohort studies are having to reconsider their position on this issue.

An advisory report was commissioned from Dr Susan Wallace in 2011 (Appendix L), and extensive discussion has followed on from this report.

At present, the situation remains that no feedback about genotypes is returned to individual study participants. However, the current version of the award letter contains a number of stipulations (see section 4.8) of which the sixth is detailed below. All potential users are therefore aware of the ACCC's concern about this increasingly important issue, and the attention of any applicant whose proposal appears to have a particular risk of generating actionable clinical findings is specifically referred to that stipulation. For example, from one of the approvals in February 2014: *in approving*

your request the Committee asked that your attention is drawn to point 6 of the stipulations below which relate to incidental findings of clinical significance….

Until any setting arises in which the current position is viewed as being untenable our current feedback strategy will be maintained. If that *does* occur we may necessarily have to refer to the Ethics Committee for advice. In the longer run, particularly if another Biomedical Survey is undertaken, a more definitive solution may become possible, for example, a more flexible consent may be used for genomic material collected in that second sweep. In the meanwhile, we will ensure that the study's feedback policy remains consistent with any proposals or strategies recommended by the funders.

6. Incidental findings of clinical significance and potential benefit

In signing their original consent forms for inclusion in the 1958BC Biomedical Survey (2002-2003), consenting participants agreed that they would not receive feedback about any individual genetic results: "...no information found in the DNA will be given to me" (NCDS Medical Follow-Up, Consent Form 2 – blood samples). In keeping with this wording the current policy of the ACCC is that no genotypic information (regardless of its nature) will be returned to cohort members.

To date, most informed commentators have seen this position as 'good practice' because nobody has really known how to interpret the clinical relevance of the genetic variants that have been identified: their effects have typically been rather small and there has been no agreed way in which to respond to the limited increases in risk they may convey. But in common with many of the world's major cohort studies and biobanks, the 1958BC recognises that national and international views of what constitutes 'best practice' might be about to change. For example, as outlined by a senior international commentator in the field[2], it is possible that in the future it may become mandatory to report genetic results to participants if they satisfy three key requirements:

***(i) scientific validity** (the genotyping is of adequate quality);*

***(ii) clinical significance** (the disease or condition caused by the genetic variant is potentially serious) , and*

***(iii) potential benefit** (i.e. a valid approach exists to prevent or cure the condition/disease of concern and that early knowledge of the genetic risk to which an individual is exposed could enhance the efficacy of that prevention/cure).*

At present a change in what is seen as best practice remains no more than a hypothetical possibility, but findings that satisfy the three stated criteria are likely to become more common as the global scientific focus moves to full sequencing of genes and/or longer segments of DNA. The ACCC therefore wishes to help contribute to the national and international evidence-base on which any future strategic decisions might be made regarding policy for feeding back genetic results.

For this reason, ***the ACCC now requires that if in the course of any analysis of DNA from any participant in the 1958BC, a genetic variant is found that could potentially be viewed as meeting all three of the criteria stated above, that information must be transmitted to the ACCC.***

At this stage this is no more than an exercise in collection of key data to assist us in developing an appropriate future strategy for the 1958BC – transmission of any information in this manner does not absolve the research group which generates the relevant finding from having their own internal policy to deal with this globally recognised problem. It is also important to ensure that your research group policy is consistent with the facts that: (1) at present NO genetic information can be returned to 1958BC participants; and (2) even if that policy were to change, all such contacts with cohort members would necessarily be undertaken by the Centre for Longitudinal Studies (contactable via ACCC). These requirements are immutable under any

circumstances – even at the direction of an ethics committee that has reviewed your (the research group's) project.

4.8 Stipulations on which applicants sign off to on the application form or are included in the award letter when it is sent to successful applicants

1. Data and samples from the 1958BC resource cannot be used for commercial purposes and any commercial involvement would breach the basis on which the access has been awarded.

2. Third party sharing of either data or biosamples is strictly prohibited. Any third party seeking to use the data, samples or derived variables or genotypes must apply directly to the Access Committee for CLS Cohorts to obtain access permission in their own right.

3. The Access Committee requires that, where possible, individual level data items created de novo are made available to other users in accordance with contemporary best practice and taking appropriate account of ethico-legal restrictions and recognising any potential risks of disclosures of summary level genotypes . If you believe that there is some reason that you can't meet this stipulation, please contact the Secretariat for the Access Committee.

4. For applications involving linked phenotype and genotype data it is important to note that once an award has been made, any future additions to the dataset (for example, if an additional linked phenotype variable is required) will have to be processed by the 1958 Birth Cohort Access Committee (Technical Review Team) and must comply with the original application. If you do need additional variables to be added, you should therefore inform the Secretariat of the Access Committee.

5. Applicants are reminded that the Terms and Conditions for the cohort explicitly forbid any attempt to identify individuals or to compromise or otherwise infringe the confidentiality of information on data subjects and their right to privacy.

6. Incidental findings of clinical significance and potential benefit

In signing their original consent forms for inclusion in the 1958BC Biomedical Survey (2002-2003), consenting participants agreed that they would not receive feedback about any individual genetic results: "...no information found in the DNA will be given to me" (NCDS Medical Follow-Up, Consent Form 2 – blood samples). In keeping with this wording the current policy of the ACCC is that no genotypic information (regardless of its nature) will be returned to cohort members.

To date, most informed commentators have seen this position as 'good practice' because nobody has really known how to interpret the clinical relevance of the genetic variants that have been identified: their effects have typically been rather small and there has been no agreed way in which to respond to the limited increases in risk they may convey. But in common with many of the world's major cohort studies and biobanks, the 1958BC recognises that national and international views of what constitutes 'best practice' might be about to change. For example, as outlined by a senior international commentator in the field2, it is possible that in the future it may become mandatory to report genetic results to participants if they satisfy three key requirements:

(i) scientific validity (the genotyping is of adequate quality);

(ii) clinical significance (the disease or condition caused by the genetic variant is potentially serious) , and

(iii) potential benefit (i.e. a valid approach exists to prevent or cure the condition/disease of concern and that early knowledge of the genetic risk to which an individual is exposed could enhance the efficacy of that prevention/cure).

At present a change in what is seen as best practice remains no more than a hypothetical possibility, but findings that satisfy the three stated criteria are likely to become more common as the global scientific focus moves to full sequencing of genes and/or longer segments of DNA. The ACCC therefore wishes to help contribute to the national and international evidence-base on which any future strategic decisions might be made regarding policy for feeding back genetic results.

For this reason, the ACCC now requires that if in the course of any analysis of DNA from any participant in the 1958BC, a genetic variant is found that could potentially be viewed as meeting all three of the criteria stated above, that information must be transmitted to the ACCC.

At this stage this is no more than an exercise in collection of key data to assist us in developing an appropriate future strategy for the 1958BC – transmission of any information in this manner does not absolve the research group which generates the relevant finding from having their own internal policy to deal with this globally recognised problem. It is also important to ensure that your research group policy is consistent with the facts that: (1) at present NO genetic information can be returned to 1958BC participants; and (2) even if that policy were to change, all such contacts with cohort members would necessarily be undertaken by the Centre for Longitudinal Studies (contactable via ACCC). These requirements are immutable under any circumstances – even at the direction of an ethics committee that has reviewed your (the research group's) project.

REFERENCES

[1] Power C, Elliott J. Cohort profile: 1958 British birth cohort (National Child Development Study). *Int J Epid* 2006:35(1):34-41.

[2] http://www.gmc-uk.org/guidance/current/library/confidentiality.asp#Disclosures%20to%20courts%20or%20in%20connection%20with%20litigation paragraph 19

[3]Public Population Project in Genomics (P3G). Future Proofing Population Genomics. *PLoS Genetics* 2009. In press.

[4]Homer N, Szelinger S, Redman M, Duggan D, Tembe W, et al. (2008) Resolving Individuals Contributing Trace Amounts of DNA to Highly Complex Mixtures Using High-Density SNP Genotyping Microarrays*. PLoS Genetics* 4(8): e1000167. doi:10.1371/journal.pgen.1000167

[5]Gymrek M, McGuire AL, Golan D, Halperin E, Erlich Y. Identifying Personal Genomes by Surname Inference. *Science* 2013; **339**: 321-4.

APPENDICES:

APPENDIX A Access Committee for CLS Cohorts Membership

APPENDIX B Terms of Reference 2014

APPENDIX C CLS SAB Membership

APPENDIX D End User Licence

APPENDIX E UKDA Special Licence

APPENDIX F Application form

APPENDIX G 1958 Biosamples Strategy Guidelines

APPENDIX H WTCCC DAC Membership

APPENDIX I 1958 Material Transfer Agreement

APPENDIX J WTCCC Data Access Agreement

APPENDIX K 1958 Biomedical Survey Consent Forms

APPENDIX L Consent Briefing Paper

APPENDIX A

Membership of the Access Committee for CLS Cohorts

Members	Term of office
Professor Paul Burton (Chair) School of Social and Community Medicine Faculty of Medicine and Dentistry University of Bristol Bristol BS8 2BN	2009 -
Professor Blair Smith (Deputy Chair) Division of Population Health Science Mackenzie Building Kirsty Semple Way Ninewells Hospital and Medical School Dundee DD2 4RB	2009 -
Professor Barbara Maughan King's College Institute of Psychiatry Denmark Hill London SE5 8AF	2009 -
Mr Paul Bradshaw ScotCen Social Research 73 Lothian Road Edinburgh EH3 9AW	2012 (Jan) -
Professor Alissa Goodman Centre for Longitudinal Studies Institute of Education 20 Bedford Way London WC1H 0AL	2014
Professor Melinda Mills Department of Sociology University of Oxford, Manor Road, Oxford OX1 3UQ Nuffield College, New Road, Oxford OX1 1NF	2014
Dr Cathie Sudlow Division of Clinical Neurosciences, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU	2014
Professor Meena Kumari Institute for Social and Economic Research, University of Essex, Colchester, CO4 3SQ	2014

Technical Advice Team	
Dr Susan Ring School of Social and Community Medicine University of Bristol Bristol BS8 2BN	2009 -
Mr Jon Johnson Centre for Longitudinal Studies Institute of Education 20 Bedford Way London WC1H 0AL	2009 -
Mr Neil Walker Cambridge Institute of Medical Research WT/MRC Building, University of Cambridge Addenbrookes Hospital Cambridge CB2 0XY	2009 -
Dr Masssimo Mangino (UK Twins Study) Department of Twins Research & Epidemiology St Thomas's Hospital Campus Westminster Bridge Road London SE1 7EH	2011 -
Observers	
Dr Katie Finch MRC Headquarters 1 Kemble Street London WC2B 4AN	
Dr Nidhee Jadeja Wellcome Trust HQ 215 Euston Road London NW1 2BE	
Ms Rachel Hall Economic and Social Research Council Polaris House North Star Avenue Swindon SN2 1UJ	
Secretariat	
Mrs Janet Jones Department of Health Sciences University of Leicester 22-28 Princess Road West Leicester LE1 6TP	

Appendix B

Access Committee for CLS Cohorts

Terms of Reference

Scope of the Committee:

The role of the Access Committee is to take decisions on applications requesting access to electronic data and biological samples from the CLS birth cohort studies[1]. The aim is to allow important research to proceed while minimizing risks[2]. The Committee is not concerned directly with the scientific merit of proposals but, rather, it addresses risks and benefits in determining whether access should be granted. However, in some instances, particularly when access is requested to a finite resource, judging the balance between risks and benefits may demand that appropriate consideration be taken of the scientific merit of the proposal. The Committee also provides a source of strategic advice and support to the funders.

As a distinct, though related, role, the ACCC is also the Committee of Final Appeal for any disputed decisions relating to access applications to the UK Twins study.

Terms of Reference:

- To establish policies and procedures for applications to access all types of data collected through the Birth Cohort Studies. This includes mechanisms by which responsibility for taking decisions on access to non-sensitive data could rest with the Data Custodians[3] of the birth cohort studies and procedures whereby low risk data are available to researchers through designated archival services operating to approved standards of access and security.

- To consider and authorise individual applications requesting access to electronic data and/or biological samples from the birth cohort studies (where responsibility has not previously been designated to the Data Custodians). A framework of precedents will be established to ensure timely consideration of subsequent requests and, where appropriate, allowing Data Custodians to approve specific classes of applications.

- To address the following issues in determining whether access should be granted:
    - Medical ethics
    - Sensitivity of data
    - Statistical disclosure
    - Governance of the data – what consent has been given by cohort members and/or data owners
    - General risks – public perception, risk to continuation of cohort, sample depletion.
    - Confidentiality of data

- If appropriate, to take advantage of third party specialist knowledge, particularly where an application has not already been through established peer review mechanism.

- To provide strategic advice to funders to help them develop and maintain efficient and effective data, sample and tissue access mechanisms both nationally and internationally. This advice may include direct input as a multi-disciplinary group of experts as well as the identification of specific issues needing special consideration by funders. In particular, the Committee will advise the funders regarding the facilitation of stream-lined access of research users to data and samples while simultaneously respecting and securing the rights and well-being of study participants and of the cohort studies themselves. In providing such advice the Committee will give appropriate recognition to the extensive input of individual scientists and research groups to the development and maintenance of these important national studies.

---

[1] Currently: NCDS (1958 Birth Cohort) – including data collected, and physical samples generated, through the biomedical sweep, BCS70 (1970 Birth Cohort) and Millennium Cohort Study (2000/01), Next Steps (formerly known as the Longitudinal Study of Young People in England, LSYPE)

[2] See for example Richard Thomas and Mark Walport "Data Sharing Review Report" July 2008 , page 70: http://www.justice.gov.uk/docs/data-sharing-review-report.pdf

[3] Data Custodians are representatives of the PI team responsible for managing the birth cohort studies

- The Committee is not the final decision-making group for substantive issues of strategy or policy. The CLS Strategic Advisory Board (SAB) will provide strategic advice to CLS on policies for use of access to the resource. This advice will be given in light of funder policies and, in particular, guidance from the Expert Advisory Group on Data Access (EAGDA). The CLS director will be responsible for acting on this advice. The Access Committee may bring key issues to the notice of the SAB, via its chair who will sit *ex officio* on the SAB. It may also submit operational or strategic documents for consideration by the SAB, but the latter will maintain ultimate decisional power.

- To provide a Committee of Final Appeal for any disputed decisions relating to access applications to the UK Twins study

- The CLS SAB will ratify the Terms of Reference for the ACCC and will act as the appeals body for disputed decisions relating to the CLS cohorts (but not for decisions relating to UK Twins).

Audit Procedures:

The Committee will maintain an effective audit mechanism. All applications will be reviewed, and progress recorded and evaluated. The audit mechanism will include oversight of returned data.

Membership:

- Independent Chair – who will sit on the CLS Strategic Advisory Board
- Independent Deputy Chair – Voting member of the Committee appointed by the Chair
- Up to five (5) additional independent members (representing a range of medical/genetic and social science experts that are all capable of enacting the roles identified above)
- A representative from CLS will form part of the Committee as a voting member
- Tenure of membership – 3 years, with option to extend for a further 3 years after the first term only. Appointment to the Committee will be staggered in order to ensure continuity of membership.

Quoracy arrangements:

The full Committee involves three groups at its meetings: independent full members; technical review team members; and invited observers including representatives from each of the three funders (MRC, WT and ESRC) and the Principal Investigator of the main 1958BC grant. Specialist technical staff (members of the technical review team) includes Senior Officers from: (1) the Institute of Education (CLS); (2) Cambridge Institute for Medical Research; (3) ALSPAC laboratories; and (4) UK Twins Study

- Members of the technical review team are strongly encouraged to attend all meetings.
- Quoracy formally requires: "the attendance of three full independent members (with at least one independent member with genetic expertise and one with social science expertise) and that either the Chairman or the Deputy Chair must be present for continuity". For face to face meetings, where it is unavoidable, attendance of a member by teleconference, will count as being present.

Modus operandi:

The committee will judge applications using the criteria and protocols outlined in a document entitled "*Policy for use and oversight of samples and data arising from the 1958 Birth Cohort"* (Annex A). This document may change from time to time. The current version is dated 9th August 2010 and the designated assessment criteria are as follows:

Assessment criteria

- Has the application been submitted by *bona fide* researchers?
- Does the application violate (or potentially violate) any of the ethical permissions granted to the study or any of the consent forms signed by the participants or their guardians?
- Does the application run a significant risk of upsetting or alienating cohort members or of reducing their willingness to remain as active participants of the particular CLS cohort of which they are a participant?
- Does the application address topics that fall within the acknowledged remit of the cohort in which they are a member, as understood by participants?

- Does the application request access to an infinite resource (data or cell line DNA) or a finite resource ?
- If the request is for a finite resource, then the application is seen as being in competition with other potential applicants (both current and future) and the quality of the science is reviewed formally (if necessary using independent external reviewers). Successful peer review by a major funder will usually be taken as evidence of appropriate quality.

Decision making:

Decisions of the Committee, including whether to grant access to a particular application, will generally be by consensus. In the unusual event of the Committee being unable to reach a clear consensus (as judged by one or more independent members) the decision will be put to a vote by the Chair, Deputy-Chair and independent members of the committee. If the votes are split evenly, the chairman will have a casting vote.

Frequency of Meetings:

The Committee will meet quarterly face to face, and on a monthly basis by teleconference.

UK Twins:

UK Twins has its own primary access committee and generates very little work for ACCC. In the three years since ACCC was invited - via Wellcome Trust - to take on the role of Committee of Final Appeal for UK Twins, no appeals have been referred up to us. Critically, if and when an appeal is submitted, the ACCC will be decisional. Because UK Twins is not a CLS cohort, ACCC will not refer issues generated by UK Twins up to CLS SAB.

Access Committee for CLS Cohorts

Committee Membership

Members	Term of office
Professor Paul Burton (Chair) School of Social and Community Medicine Faculty of Medicine and Dentistry University of Bristol Bristol BS8 2BN	2009
Professor Blair Smith (Deputy Chair) Division of Population Health Science Mackenzie Building Kirsty Semple Way Ninewells Hospital and Medical School Dundee DD2 4RB	2009
Professor Barbara Maughan King's College Institute of Psychiatry Denmark Hill London SE5 8AF	2009
Mr Paul Bradshaw ScotCen Social Research 73 Lothian Road Edinburgh EH3 9AW	2012
Professor Alissa Goodman Centre for Longitudinal Studies Institute of Education 20 Bedford Way London WC1H 0AL	2014
Professor Meena Kumari Institute for Social and Economic Research, University of Essex, Colchester, CO4 3SQ	2014
Professor Melinda Mills Department of Sociology University of Oxford, Manor Road, Oxford OX1 3UQ Nuffield College, New Road, Oxford OX1 1NF	2014
Dr Cathie Sudlow Division of Clinical Neurosciences, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU	2014

Technical Advice Group	
Dr Susan Ring School of Social and Community Medicine University of Bristol Bristol BS8 2BN	2009 -
Mr Jon Johnson Centre for Longitudinal Studies Institute of Education 20 Bedford Way London WC1H 0AL	2009 -
Mr Neil Walker Cambridge Institute of Medical Research WT/MRC Building, University of Cambridge Addenbrookes Hospital Cambridge CB2 0XY	2009 -
Dr Masssimo Mangino (UK Twins Study) Department of Twins Research & Epidemiology St Thomas's Hospital Campus Westminster Bridge Road London SE1 7EH	2011 -
Observers	
Dr Katie Finch MRC Headquarters 1 Kemble Street London WC2B 4AN	
Dr Nidhee Jadeja Wellcome Trust HQ 215 Euston Road London NW1 2BE	
Ms Rachel Hall Economic and Social Research Council Polaris House North Star Avenue Swindon SN2 1UJ	
Secretariat	
Mrs Janet Jones Department of Health Sciences University of Leicester 22-28 Princess Road West Leicester LE1 6TP	

APPENDIX C

The Centre for Longitudinal Studies Strategic Advisory Board (SAB) Membership

Members
Mr Richard Bartholomew (Chair)
Ms Launa Anderson
Professor Paul Burton (ex officio)
Dr Pamela Davis-Kean
Dr Katie Finch (MRC)
Professor Andy Furlong
Ms Rachel Hall (ESRC)
Ms Lisa Harker
Professor Kath Kiernan (Observer)
Professor Chris Taylor
Dr Jimmy Whitworth
Dr Janet Valentine (MRC)
Professor Neil Pearce

Ex officio
Professor Jane Elliott
Professor Emla Fitzsimons
Professor Alissa Goodman
Dr Alice Sullivan

ESRC Office
Michelle Dodson
Julie Dunsby (Minutes)

APPENDIX D

END USER LICENCE

PUBLIC VERSION
14 FEBRUARY 2013
Version: 04.00

T +44 (0)1206 872572
E susan@essex.ac.uk
www.data-archive.ac.uk

UK DATA ARCHIVE
UNIVERSITY OF ESSEX
WIVENHOE PARK
COLCHESTER
ESSEX, CO4 3SQ

WE ARE SUPPORTED BY THE **UNIVERSITY OF ESSEX**, THE **ECONOMIC AND SOCIAL RESEARCH COUNCIL**, AND THE **JOINT INFORMATION SYSTEMS COMMITTEE**

Contents

1. End User Licence (EUL) text

This Agreement is made between you and the University of Essex (also referred to as the "registrar") and the service funders in order to provide you (the "End User") with the right to use the collections provided via the UK Data Service, according to the terms below.

In this agreement:

"**Data Team**" means in relation to a particular data collection, the registrar, the relevant data service providers, and (to the extent that the Special Conditions and/or metadata specific to a particular data collection expressly provide) the service funders, data collection funders and/or original data creators or depositors.

"**data service provider**" means the persons or organisations that directly provide you with the data collections (on behalf of the service funder). The data service provider for a particular data collection is identified in the Special Conditions and/or metadata applicable to that data collection;

"**service funder**" means the persons or organisations that fund the data service provider as defined above. The service funder for a particular data collection is identified in the Special Conditions and/or metadata applicable to that data collection;

"**data collection funder**" means the persons or organisations that funded the collection and/or creation of the data collections. The data collection funder for a particular data collection is identified in the Special Conditions and/or metadata applicable to that data collection;

"**original data creator or depositor**" means the persons or organisations that originally collected, created or deposited the materials making up the data collections and/or who own the intellectual property rights in the data collections. The original data creator or depositor for a particular data collection is identified in the Special Conditions and/or metadata applicable to that data collection;

"**registrar**" means the person or organisation responsible for the system that registers End Users and issues them with End User Licences (being the University of Essex);

"**Special Conditions**" means any further conditions applicable to the use of one or more data collections by an End User, as notified to the End User in accordance with paragraph 5 of the End User Licence;

"**metadata**" means any additional or bibliographic information about one or more of the data collections, as notified to the End User from time to time. Metadata may be supplied by electronic means.

I (the "End User") agree to the following conditions of use in consideration of the data collections being made available to me through the various contributions of each member of the Data Team:

1. To use the data collections only in accordance with this End User Licence and to notify promptly the registrar and the data service provider of any breach of its terms in writing or of any infringements of the data collections of which I become aware.

2. To use and to make personal copies of any part of the data collections only for the purposes of not-for-profit research or teaching or personal educational development. To obtain permission prior to using part or all of the data collections for commercial purposes by contacting the registrar and/or relevant data service provider, where relevant, in order to obtain an appropriate licence from the rights holder(s) in question or their permitted licensee if one is available.

3. That this Licence does not operate to transfer any interest in intellectual property from the data collection funders, service funder(s), the data service providers, the original data creators, producers, depositors, copyright or other right holders (including without limitation the ONS or the Crown) to me. That any rights subsisting in materials derived now or in the future from the data collections which are the intellectual property of the Crown are hereby assigned (by way of assignment of present and future intellectual property) to the Crown by this Licence to the extent not already vested in the Crown. To take all steps necessary to give effect to this Clause (including by executing further written documentation).

4. That the Licence and the data collections are provided by the Data Team on an "as is" basis and without warranty or liability of any kind. Any representations or warranties given by any member of the Data Team relating to this licence, expressed or implied, are excluded to the maximum extent permitted by law.

5. To abide by any further conditions notified to me from time to time by the registrar or the relevant data service provider that may apply to the access to, or use of, specific materials within the data collections or particular data collections. Notice of further conditions under this paragraph may be given to me by electronic means, for example, by way of a pop-up window upon my ordering one or more data collections. My acceptance of the further conditions shall be required before I gain access to the data collections in question. In this Agreement such further conditions are referred to as Special Conditions.

6. To give access to the data collections, in whole or in part, or any material derived from the data collections, only to registered End Users who have entered into an End User Licence and accepted the relevant Special Conditions necessary to access and use the data collections (with the exception of data collections or material derived from data collections supplied for the stated purpose of teaching or included in publications made for the purposes set out in paragraph 2).

7. To ensure that the means of access to the data (such as passwords) are kept secure and not disclosed to a third party except by special written permission or licence obtained from the original data service provider.

8. To preserve at all times the confidentiality of information pertaining to individuals and/or households in the data collections where the information is not in the public domain. Not to use the data to attempt to obtain or derive information relating specifically to an identifiable individual or household, nor to claim to have obtained or derived such information. In addition, to preserve the confidentiality of information about, or supplied by, organisations recorded in the data collections. This includes the use or attempt to use the data collections to compromise or otherwise infringe the confidentiality of individuals, households or organisations.

9. To acknowledge, in any publication, whether printed, electronic or broadcast, based wholly or in part on the data collections, the original data creators, depositors or copyright holders, the service funders and the data service provider(s) in the form specified on the data distribution notes or in accompanying metadata received with the dataset or notified to me and without prejudice to paragraph 5 above to comply with any restrictions on my use of the data collections referred to or referenced therein or otherwise notified to me from time to time. To cite, in any publication, whether printed, electronic or broadcast, based wholly or in part on the data collections, the data collections used in the form specified on the data distribution notes or in accompanying metadata received with the dataset or notified to me.

10. To supply the relevant data service provider with the bibliographic details of any published work based wholly or in part on the data collections.

11. That the members of the Data Team may hold and process any personal data submitted by me for validation and statistical purposes, and for the purposes of the management of the service or for any other lawful purpose notified to me and to which I have consented under this Agreement in relation to a particular data collection, and they may also pass the information on to other parties such as: (i) depositors and distributors of material contained in or accessed via the data service provider; (ii) copyright and other intellectual property rights owners whose material is held by the data service

provider; as well as (iii) each member of the Data Team's organisation and (iv) my own institution or organisation, in compliance with the Data Protection Act 1998.

12. To notify the data service provider of any errors discovered in the data collections.

13. That any personal data submitted by me is accurate to the best of my knowledge, and that any changes in that personal data, including my educational or employment status, will be made known to the registrar at the earliest possible opportunity.

14. To meet any charges that may from time to time be levied by any member of the Data Team for the supply of the data collections including, where relevant, annual service fees and royalty fees.

15. At the conclusion of my research (or if earlier at any time at the request of a member of the Data Team), to offer for deposit in the data collection(s) on a suitable medium and at my own expense any new data collections which have been derived from the materials supplied or which have been created by the combination of the data supplied with other data. The deposit of the derived data collection(s) will include sufficient explanatory documentation to enable the new data collection(s) to be accessible to others.

16. I understand that breach of any of the provisions of this Agreement will lead to immediate termination of my access to all services provided by the Data Team either permanently or temporarily, at the discretion of a member of the Data Team, and may result in legal action being taken against me. I understand that where there is no breach of this Licence, it may be terminated, or its terms altered, by a member of the Data Team either after 30 days notice; or, if a service charge has been paid in advance, at the end of the period for which payment has been made, whichever is the longer. The failure to exercise or delay in exercising a right or remedy provided by this Agreement or by law does not constitute a waiver of the right or remedy or a waiver of other rights or remedies.

DISCLAIMERS

To the extent that applicable law permits:

a. The members of the Data Team bear no legal responsibility for the accuracy or comprehensiveness of the data supplied.

b. The members of the Data Team accept no liability for, and I will not be entitled to claim against them in respect of, any direct, indirect, consequential or incidental damages or losses arising from use of the data collections, or from the unavailability of, or break in access to, the service, for whatever reason.

c. Whilst steps have been taken to ensure all licences, authorisation and permissions required for the granting of this Licence have been obtained, this may not have been possible in all cases, and no warranties or assurance are given in this regard. To the extent that additional licences, authorisations and permissions are required to use the data collections in accordance with this Licence, it is the End User's responsibility to obtain them.

d. I agree to indemnify and shall keep indemnified each member of the Data Team against any costs, actions, claims, demands, liabilities, expenses, damages or losses (including without limitation consequential losses and loss of profit, and all interest, penalties and legal and other professional costs and expenses) arising from or in connection with any third party claim made against any member of the Data Team relating to my use of the data collections or any other activities in relation to the data where such use is in breach of this licence.

If the whole or any part of a provision of this Agreement is void, unenforceable or illegal for any reason, that provision will be severed and the remainder of the provisions of this Agreement will continue in full force and effect as if this Agreement had been executed with the invalid provision eliminated.

This Agreement may be enforced separately in relation to each data collection provided to the End User by any member of the Data Team and the End User. No other persons may enforce this Agreement under the Contract (Rights of Third Parties) Act 1999.

This Agreement (which is the entire agreement between the parties and supersedes any previous agreement between them) may be varied in writing by agreement of the relevant service funders, the registrar, and the End User (who may give its consent to such variations by electronic means). No consent from any other party is required to vary or rescind this Agreement.

This Agreement and any documents to be entered into pursuant to it shall be governed by and construed in accordance with the laws of England and Wales and each Party irrevocably submits to the exclusive jurisdiction of the courts of England and Wales over any claim or matter arising under or in connection with this Agreement and the documents entered into pursuant to it.

2. End User Licence (EUL) summary text

Seventeen points to help you understand the End User Licence (EUL). These pointers are for general guidance and you must read and understand the full EUL before agreeing to it. By accepting the EUL, you agree:

1. to use the data in accordance with the EUL and to notify the UK Data Service of any breach you are aware of
2. not to use the data for commercial purposes without obtaining permission and, where relevant, an appropriate licence if commercial use of the data is required
3. that the EUL does not transfer any interest in intellectual property to you
4. that the EUL and data collections are provided without warranty or liability of any kind
5. to abide by any further conditions notified to you
6. to give access to the data collections only to registered users (who have accepted the terms and conditions, including any relevant further conditions). There are some exceptions relating to teaching.
7. to ensure that the means of access to the data (such as passwords) are kept secure and not disclosed to anyone else
8. to preserve the confidentiality of, and not attempt to identify, individuals, households or organisations in the data
9. to use the correct methods of citation and acknowledgement in publications
10. to send the UK Data Service bibliographic details of any published work based on our data collections
11. that personal data about you may be held for validation and statistical purposes and to manage the service, and that these data may be passed on to other parties
12. to notify the UK Data Service of any errors discovered in the data collections
13. that personal data submitted by you are accurate to the best of your knowledge and kept up to date by you
14. to meet any charges that may apply
15. to offer for deposit any new data collections which have been derived from the materials supplied
16. that any breach of the EUL will lead to immediate termination of your access to the services and could result in legal action against you

APPENDIX E

SPECIAL LICENCE:

Centre for Longitudinal Studies

PUBLIC VERSION
18 MAY 2011

T +44 (0)1206 872572
E susan@data-archive.ac.uk
www.data-archive.ac.uk

UK DATA ARCHIVE
UNIVERSITY OF ESSEX
WIVENHOE PARK
COLCHESTER
ESSEX, CO4 3SQ

WE ARE SUPPORTED BY THE **UNIVERSITY OF ESSEX**, THE **ECONOMIC AND SOCIAL RESEARCH COUNCIL**, AND THE **JOINT INFORMATION SYSTEMS COMMITTEE**

Special Licence – Centre for Longitudinal Studies

Definitions

- Licence holder – the principal licence holder and associated parties to this licence specified in sections 1, 2, and 4
- Data depositor – Centre for Longitudinal Studies
- Data – the collections detailed in section 8.2 of this licence
- Dispute arbitrator – ESRC

The data to which this Licence, known as a 'Special Licence', permits access are those of the data depositor and are held under 'Special Conditions', as specified in section 5 of the Economic and Social Data Service (ESDS) End User Licence (EUL).

This Special Licence specifies the conditions for access for statistical research purposes, the obligations of the researcher/s and the measures for protecting and respecting the confidentiality of statistical data.

The Special Licence grants the licence holder access solely for the purposes specified. The licence holder

- will take all necessary administrative, technical and organisational measures to ensure that the data are used only in the manner stated and for the research purposes specified
- will not process, disseminate or otherwise allow any of the data to be made available or used for any other purpose whatsoever and will remain bound by this obligation even after expiry or termination of the contract
- will not attempt to use these data after the expiry of the Licence
- will guarantee that none of these data are distributed to third parties
- will not attempt to identify by any means whatsoever, any individual statistical unit, nor will the licence holder claim to have done so
- will apply methods and standards specified in this licence for disclosure control for any outputs

Acceptance by the licence holder of the further conditions specified below is required before access to the data is granted

The Licence Holder is advised to read the Completion Notes at the end of the Licence before proceeding.

1. PRINCIPAL LICENCE HOLDER

Principal licence holder's details

Name	Employing organisation	Address of organisation	Position in organisation	Tel. No.	Email

2. RESEARCH TEAM

Details of each member of the team

Name	Employing organisation	Address of organisation	Position in organisation	Tel. No.	Email

3. ESDS EUL

The licence holder / and all members of the research team (delete as applicable) has / have registered with the ESDS or the Census Registration Service (CRS) and the registration and the EUL have been accepted

☐ (Tick to confirm)

4. RESPONSIBILITY for the licence holder's use of the data

4.1 ORGANISATION with the ultimate responsibility for the licence holder

Name of organisation	Address of organisation

4.2 ORGANISATION'S REPRESENTATIVE

The person with the authority to represent the organisation

Name	Organisation	Position in organisation	Tel. No.	Email

4.3 ORGANISATION employing the licence holder (where this is different to 4.1)

Name of organisation	Address of organisation

4.4 ORGANISATION'S REPRESENTATIVE

The person with the authority to represent the employing organisation

Name	Organisation	Position in organisation	Tel. No.	Email

5. FUNDING: Details of external funding that has been sought

5.1 Organisation funding the research project

Name of organisation	Address of organisation

5.2 Funding

The licence holder confirms that funding has been sought ☐ (tick to confirm)

Funding has been obtained: YES / NO / NOT YET HEARD (please delete as applicable)

6. SITE OF ACCESS

Name of organisation	Address of organisation

7. DURATION OF ACCESS

Period of access specified must not exceed 2 years

From ☐ dd/mm/yy

To ☐ dd/mm/yy

(If it is necessary to extend the period of access, application must be made to the UK Data Archive prior to the expiry of the Licence)

8. TITLE OF RESEARCH PROJECT including UK Data Archive usage number

8.1 Where research is part of a larger programme, please give details

8.2 Title of the dataset(s) and the study number(s) to which access is required

Title of dataset	**Study number**

9. PURPOSE FOR ACCESS

9.1 Details to include:

(i) A brief summary of up to 200 words describing the aims of the study/research project

(ii) Full description of the purpose/s for which the data are requested

(iii) A justification as to why access to the special conditions version of the data is needed and why data available under the EUL is not sufficient for the purposes

9.2 A description of the analyses that will be performed on the data

10. USE OF THE DATA FOR COMMERCIAL GAIN

All signatories (other than the data depositor and the UK Data Archive) guarantee that these data will not be used for personal or commercial gain. The focus of the project is statistical research/analysis and the data will not be used for any other purpose.

[Statistics arising from the use of these data can be used for any purpose, subject to meeting the standards for disclosure control detailed in section 11]

11. PRODUCTS and PUBLICATIONS

11.1 Protecting confidentiality

The licence holder is aware that the microdata may allow individuals to be identified. Any outputs made available to anyone other than those named on the Licence, must meet the guarantee contained in the Code of Practice for Official Statistics and the Protocol on Data Access and Confidentiality, namely that no statistics are produced that are likely to identify an individual, unless specifically agreed with them.

The following rules will allow the guarantee to be kept in most cases. However, it is the responsibility of the licence holder and all signatories (other than the data depositor and UK Data Archive) to consider and protect against any other circumstances that might result in the disclosure of the identity of an individual.

11.2 Disclosure Protection

The licence holder will apply the supplied methods and standards below for disclosure control for any outputs released beyond the research team.

The licence holder will avoid small sample base numbers because they will be unreliable. For example, percentages based on small counts will have very wide confidence intervals.

Supplied methods and standards:

(i) Tables that contain very small sample numbers in some cells may be disclosive. The licence holder will ensure that tables do not report numbers or percentages in cells based on only 1 or 2 cases. Cells based on 1 or 2 cases should be combined with other cells or, where this is not appropriate, reported as 0 percent.

(ii) The licence holder will ensure that all tables report weighted values, where weights are available.

(iii) Tables and other outputs derived from data accessed through a Special Licence will not be published in a form where the level of geography would threaten the confidentiality of the data. Typically, outputs with a geography of region or greater can be considered safe.

Outputs with a geography between Local Authority and region can in some circumstances introduce disclosure risk. Where there is any doubt, the licence holder must contact the UK Data Archive to gain confirmation of the confidentiality of any outputs for publication with geography below region.

No outputs will be published with a geography below local authority.

(iv) Although most outputs from models or other statistical analysis will not be disclosive, the licence holder will ensure that individuals, households or organisations cannot be identified. In particular, results based on very small numbers should be avoided. Any result that refers to unit records, e.g. a maximum or minimum value should not be published. Models should not report actual values for residuals

(v) Graphical outputs should be based on non-disclosive data. The licence holder will take particular care not to report extreme outliers.

11.3 Intended outputs / publications arising from the use of these data

11.3.1 The data depositor reserves the right to comment on statistical issues raised by publications and to scrutinise outputs before publication for disclosure control purposes. Where the data depositor so requires, the licence holder must supply the data depositor with a copy of any proposed publication, based wholly or in part on the data collections accessed, to enable the data depositor to consider it and comment as regards compliance with the conditions for disclosure protection and for changes to be made to the publication in the light of those comments.

The licence holder will make any [reasonable] changes that are required by the data depositor in order to make the proposed publication comply with these conditions.

11.3.2 The licence holder must supply to the UK Data Archive the bibliographic details of any published work based wholly or in part on the data collection/s accessed. Details are to be provided on publication.

12. MINIMUM INFORMATION REQUIRED

The licence holder confirms that access to the data is required in order to meet the aims of the project and that the access is proportionate and not excessive to the stated statistical purpose.

13. MATCHING or LINKING

Under this Licence, it is forbidden to match or attempt to match individual or household records to any other data source at the level of individual or household. Only area-level descriptors or other group-level classifications may be matched for analysis purposes.

14. DUPLICATION

The licence holder agrees that:

14.1 Any intended duplication of the data will only be for the purpose of making personal copies to aid their own research and analysis

14.2 No duplication of the data for any other purpose may take place.

15. EXPIRY OF ACCESS PERIOD

15.1 At the end of the access period, the licence holder agrees to destroy all copies of the data, including temporary copies, CDs, printed copies, personal copies, back-ups, derived datasets and all electronic copies.

15.2 The licence holder will ensure that the data are destroyed to the standards specified in the document *Microdata Handling and Security: Guide to Good Practice* (link attached)

15.3 After expiry of this Licence, the licence holder will sign and send to the UK Data Archive, a declaration to confirm that all copies of the data have been destroyed and to the required standards, or that the data

have been returned to the UK Data Archive for destruction.

16. SECURITY OF THE DATA

The licence holder guarantees to preserve at all times the confidentiality requirements associated with the data and to meet the conditions specified in the EUL. Wrongful disclosure will attract penalties as detailed in section 17 below and outlined in the document *Microdata Handling and Security: Guide to Good Practice*.

Confidentiality requirements:

The licence holder will ensure that:

16.1 Access to the data, any copies made of the data and the information contained in them is limited solely to the person who has signed this Licence and the research team, who have also signed the Special Licence.

16.2 The confidentiality of the data will be preserved in outputs and publications, as detailed in section 11.

16.3 The means of access to the data (such as passwords or pass-phrases) are kept secure and not disclosed by the Licence Holder or any member of the research team to any other individual, under any circumstances.

16.4 Data will only be accessed, in an institutional setting, via a stand-alone PC or a closely controlled LAN with restricted access. Access to the PC or LAN will be via password or pass-phrase.

16.5 Hard copies and backups of data are to be stored in a secure, access restricted filing cabinet

16.6 Stand-alone PCs and LANs, which have Internet access via broadband connection (and not through a secure organisational provider, e.g. JANET), will not have live Internet links while the data are in clear/unencrypted text on the machine. At such times the Internet will be disconnected and the broadband cable will be physically disconnected from the PC.

16.7 Stand-alone PCs and LANs, which have Internet access via dial-up telephone connection (and not through a secure organisational provider, e.g. JANET), will not have live Internet links while the data are in clear/unencrypted text on the machine.

16.8 Data requested under the Special Licence will only be accessed at a site that has security standards that meet the requirements outlined in the document *Microdata Handling and Security: Guide to Good Practice*.

16.9 Data will not be accessed at a private residence

16.10 The University of Essex and the data depositor reserve the right to conduct an on-site audit of the confidentiality and security procedures and practices for guaranteeing the security and confidentiality of the data covered by this Licence, or to require a report of such an audit.

16.10.1 For the purpose of conducting an audit, the University of Essex (or the UK Data Archive, on behalf of the University of Essex) or the data depositor may enter the premises where the data are stored and processed without notice at any reasonable time. The organisation with ultimate responsibility for the licence holder undertakes to allow the University of Essex or data depositor access for this purpose.

16.10.2 The data depositor further requires that the organisation with ultimate responsibility for the licence holder provides to the UK Data Archive, copies of any audits of these arrangements, conducted for the organisation or the licence holder, during the period of the Licence, including any audit implementation plans.

17. BREACH PROCEDURES

17.1 Any breach of any of the provisions of this Licence will result in the immediate termination of the licence holder's access to the data, the termination of the licence and the prohibition of any further access to the data depositor's data via the Special Licence. It will also lead to immediate termination of the services provided by the UK Data Archive data team, either permanently or temporarily (as stated in section 16 of the EUL).

17.2 The breach of any of the provisions of this Licence may result in sanctions being sought against the licence holder. These may include legal proceedings being taken by the data depositor for breach of obligations under statute or common law.

[Details of sanctions that may be sought can be found in the Completion Notes, section 2, 17.]

17.3 The licence holder is required to report promptly a breach of any of the terms of the Licence. Failure to disclose details is a fundamental breach of this Licence.

18. DISPUTE PROCEDURES

Any disputes arising from the use of the data and/or the terms of this licence will be resolved initially between the UK Data Archive, on behalf of the University of Essex and the principals to the agreement (the Licence holder and the organisation with ultimate responsibility for the Licence holder). Otherwise, outstanding issues will be referred to the dispute arbitrator.

19. AGREEMENT

19.1 The licence holder and, where the research project is undertaken by a research team, all members of the research team, agree/s to:

(i) comply with the terms and requirements of this Special Licence.

(ii) comply with any additional conditions that the data depositor may consider necessary before approving this Special Licence. Such conditions will be added to the Licence by the data depositor, at the time of approval, and notified to the licence holder by the UK Data Archive upon receipt of approval from the data depositor. Downloading the data by the Licence Holder will signify acceptance of such additional conditions.

(iii) continue to meet the terms of the End User Licence (EUL). Where there is disparity between the EUL and the Special Licence, the Special Licence will take precedence, unless identified explicitly in writing.

(iv) read the document *Microdata Handling and Security: Guide to Good Practice* and abide by the principles for use of the data, detailed therein.

19.2 The licence holder and, where the research project is undertaken by a research team, all members of the research team, understand/s:

(i) should circumstances require, the Licence may be terminated or suspended, access to the data terminated or suspended, or the terms of the Licence altered, by a member of the Data Team (as defined in the ESDS EUL) or by the data depositor. This may take immediate effect, or after a period of 30 days notice.

(ii) the principles of the Freedom of Information Act apply and nothing provided in this Licence is confidential to the licence holder or to the data depositor. To disclose the details of the Licence would not be a breach of any duty of confidence and therefore the details would be made available to the public on request and may be included as part of the metadata attached to any of the outputs arising

from the access.

(iii) these data are provided in good faith and, to the best of the data depositor's knowledge and ability, are free of error at the time of supply. The data depositor and the UK Data Archive will not be responsible for any errors, omissions or mistakes contained in the users' dataset nor for any consequences or liabilities arising therefrom. The data depositor's liability shall be limited to re-supply of corrected materials.

19.3 The signatories believe that the Licence is compliant with the statements of principle in the Code of Practice for Official Statistics (the Code) and the specific requirements of the Protocol on Data Access and Confidentiality (PDAC). Where this Licence may appear to contradict the statements of principle in the Code or the specific requirements of the PDAC, the Code and the PDAC take precedence, unless explicitly stated.

The Code of Practice for Official Statistics is available from:

http://www.statisticsauthority.gov.uk/assessment/code-of-practice/index.html

The National Statistics Protocol on Data Access and Confidentiality is available from:

http://www.ons.gov.uk/ons/guide-method/the-national-statistics-standard/code-of-practice/protocols/index.html

20. SIGNATURES

20.1 Licence holder and research team

Name of licence holder	Signature of licence holder	Date

Names of Research Team members	Signatures of Research Team members	Date

20.2 ORGANISATION WITH RESPONSIBILITY FOR THE LICENCE HOLDER

The .....(name of organisation) undertakes to accept ultimate responsibility for the licence holder's access to the data stated above

Name of organisation's representative	Signature of organisation's representative	Date

20.3 ORGANISATION WITH RESPONSIBILITY FOR THE LICENCE HOLDER (as employer) (complete where this is different to 20.2)

The .....(name of organisation) undertakes to accept responsibility for the licence holder's access to the data stated above, as the employing organisation.

Name of organisation's representative	Signature of organisation's representative	Date

21. APPROVAL

21.1 UK Data Archive

The UK Data Archive on behalf of the University of Essex, have screened the request and confirms that it meets the terms of the agreement between the data depositor and the University of Essex for access to these data

Name of UK Data Archive representative	Signature of UK Data Archive representative	Date

21.2 Approval of the data depositor

The data depositor confirms that the access complies with any undertaking made at the time of collection or the scope of any consent given.

The data depositor authorises the provision of access to these data to the licence holder under the terms specified in this Special Licence, including any additional conditions imposed by the data depositor, as stated below:

Additional conditions of access:

Name of representative for the data depositor	**Signature of representative for the data depositor**	**Date**

COMPLETION NOTES

1. General notes:

1. The Special Licence is to be used for access to data of the data depositor that are subject to special conditions and controlled access arrangements.

2. Approval to access the data is conditional upon the Licence Holder, any other named users and the 'responsible' organisation, agreeing to the terms and special conditions detailed in the Special Licence.

3. The data depositor retains the right of veto and may refuse access to the data requested by the Licence Holder. Such decision will be communicated to the Licence Holder by the UK Data Archive, together with the reason for the decision.

4. The Special Licence is to be completed by the Licence Holder, who will be the researcher requiring access to the data stated for a specific research purpose, for a time limited period. Where the researcher is part of a research team, the Licence Holder will be the head of the research team.

5. Parties to the Special Licence, who will be bound by the terms of the Licence, include:

   (i) Licence Holder

   (ii) Members of a research team, who must be identified and, in addition to the Licence Holder, will sign the Licence

   (iii) Organisation with the ultimate responsibility for the Licence Holder and any members of a research team (section 4 on the Licence and point 4 below)

   (iv) Employing organisation, where this is organisation is different to (iii)

(v) The data depositor

6. Signatories to the Special Licence:

   (i) Licence Holder

   (ii) All members of a research team

   (iii) Representative for the organisation with ultimate responsibility for the Licence Holder and any research team (see point 4 below)

   (iv) The representative of the employing organisation, where this is organisation is different to (iii)

   (v) Representative for the UK Data Archive

   (vi) Representative for the data depositor

7. Names/details of organisations to be included on the Special Licence, in addition to those listed in section 5 above:

   (i) Where the research is being externally funded, the name of the funding organisation

8. All information is to be given in plain English and full explanations are to be given where unfamiliar terminology is included.

9. Details provided are to be full, coherent and concise.

   Failure to provide adequate or comprehensive details will result in the Licence being returned to the applicant. This will delay the process and will also require the re-gaining of signatures to confirm the additional information provided.

10. The Licence Holder will not make any changes to the format and content of the clauses of the Special Licence. Changes will be identified, will delay the process and may result in the Special Licence being withdrawn.

2. Guidance on individual sections:

SECTION	NOTES
1 Licence holder details	(i) The organisation to be entered is the licence holder's employer (ii) The addition of address and telephone number information is not mandatory
2 Research team	Where access to the data is requested by a Research Team, the Special Licence is to be completed by the lead researcher who will be the licence holder. The name/s of the other member/s of the research team are to be entered in section 2.
3 The End User Licence (EUL)	Request for access to Special Licence data is conditional upon prior registration with ESDS and acceptance of the EUL.
4 Responsibility for the licence holder	(i) The organisation to be entered is that which has the ultimate responsibility for the licence holder's use of the data. This is not necessarily the organisation that employs the licence holder. (ii) The name to be entered is that of the person with the authority to

	represent that organisation: (Sections 4.2 & 20.2) See table below for further details * Supervised use of the data: Where the Licence Holder's use of the data is supervised, as may be the case with PhD students, the Supervisor is to be a member of the research team and their details included as requested in section 2.
5 Funding	Where the research is not subject to funding, enter 'N/A'.
6 Site of access	Special Licence data may not be accessed at a private residence. Data may only be accessed in an institutional setting, i.e. the site of the licence holder's employment, the site of the organisation with the ultimate responsibility for the licence holder, or the site of the funding or commissioning organisation.
7 Duration of access	The period of access stated should not be longer than the time required for conducting the research and producing outputs, with a maximum period of 2 years. Where it is necessary to extend the period of access, the Licence Holder should contact the UK Data Archive (the Archive), Support Services, in advance of the expiry of the period of access. Support Services will provide advice on the action to be followed.
8 The Research Project	(i) Where a research project does not have a usage number, the Licence Holder is to contact Archive's Support Services for guidance (ii) Details are to be included where a project is part of a larger programme or funded jointly by various organisations (iii) Access can only be requested to dataset/s that are currently available through the Archive catalogue and have a study number. Where the research project requires access to other data that are not in the catalogue, contact is to be made with the Archive for advice.
9 Purpose for access	Data held under the Special Licence are only to be accessed for statistical research purposes.
10 Use of the data for commercial purposes	The purpose for which the data are required must be statistical and the focus of the research, the resultant analysis. The prime focus for accessing the data must not be for the purpose of personal or commercial gain.
11 Products and Publications	(i) The licence holder agrees to ensure that disclosure control methodology, applied to outputs, is sufficient to ensure so that 'it would take a disproportionate amount of time, effort and expertise for an intruder to identify a statistical unit to others, or to reveal information about that unit not already in the public domain'. (Extract from the National Statistics Protocol on Data Access and Confidentiality) (ii) Where the data depositor requires sight of proposed outputs before publication, the data depositor will endeavour to comment and respond within one week of receipt. However, should circumstances require further discussion and investigation, the data depositor will notify the Archive with the minimum of delay and will be sensitive to licence

	holder's commitments and publication deadlines (iii) Where the Licence Holder has any doubts about maintaining the confidentiality of the data in the outputs, contact is to be made with the Archive User Support Services, who will contact the data depositor for advice and guidance.
12 Minimum information required	Access to the data must be proportionate to the stated statistical purpose. As part of the approval process, the data depositor requires this assurance and will take into account the researcher/organisation benefiting from the access, the type of information being accessed, the method of access, the researcher's needs and the purpose for the research.
13 Matching or Linking	Where the Licence holder wishes to conduct a matching or linking exercise that would breach the terms of the Special Licence, the Licence Holder must contact the Archive before proceeding. The Archive will contact the data depositor for a decision.
14 Duplication	(i) The Licence Holder may take personal copies of the data to assist with the specified research and analysis. However, the Licence Holder is prohibited from taking copies for any other purpose. (ii) At the end of the period of access, all copies of the data, in whatever format made, must be destroyed. See section 15 below.
15 Expiry	At the expiry of access period, the Licence Holder must agree to destroy the data and all copies made in the manner specified in the document *Microdata Handling and Security: Guide to Good Practice*.
16 Security	Data may only be accessed according to the security conditions detailed in section 16 of the Special Licence. Licence holders must note the instructions the document *Microdata Handling and Security: Guide to Good Practice* and are reminded that: (i) Data will be encrypted during transmission. However, when the licence holder accesses the data, it will appear as clear, plain unencrypted text in the format selected by the licence holder, i.e. SPSS, STATA or ASCII. At such times, the licence holder must ensure that, for PCs that have Internet access via broadband or telephone dial-up connection (and not through a secure organisational provider, e.g. JANET), the Internet is disconnected. For Broadband/internet connections, cables are to be physically disconnected from the PC. (ii) Where Internet access is through a 'secure organisational provider', it is not necessary to physically disconnect cables or disable internet systems. If there is any uncertainty as to whether an 'organisational provider' is 'secure', contact the Archive's Support Services with details of the system that is in place. (iii) Licence holders are reminded that data may not be accessed at a private residence. (iv) The data depositor and the University of Essex (or the Archive on behalf of the University of Essex) reserve the right to conduct an audit and to enter premises for this purpose. The licence holder is advised to bring this requirement to the attention of the individual with the authority to represent the organisation, before that individual and the

	licence holder sign the Licence.
17 Breach Procedures	The licence holder is reminded that a breach of any of the terms of this Licence must be reported promptly to the Archive. Failure to do so is a fundamental breach of the Licence. Sanctions that may be applied: 1. For a first offence, the penalty should be a minimum twelve-month non-discretionary suspension from access to any micro-data, applicable to the individual in question. It would generate a written warning to the institute. 2. An individual's second breach would, as a minimum, result in a suspension of access of two to five years, or permanently, on the individual, and would generate a written warning to the individual's institution. 3. If the individual has moved institutions between first and second breaches, the new institution will receive an advisory letter to include details of the 1st breach. 4. Any discretionary penalty may be decided, including permanent suspension for the individual or other staff in the relevant department, and/or pursuing in the Courts an action for breach of contract. 5. Where the breach is the result of an institution's wilful or negligent action, then a minimum penalty of a twelve-month non-discretionary suspension shall apply to the relevant department within the institution. Repeated breaches will result in a letter with discretionary penalties to the institution as a whole including suspension of all data access facilities for all the institution's staff and/or an action for breach of contract. 6. The consequences of any suspension of access (such as consequent inability to honour research contracts) will not be taken into consideration when applying minimum penalties or any of the Archive's (or, for ONS data, the National Statistician's) discretionary penalties. 7. Any appeal will be to the Archive in the first instance and may be referred to the dispute arbitrator.
18 Dispute Procedures	The Archive acts as data custodian for the data deposited at Archive by the data depositor. Therefore, where there is a dispute arising from the use of the data and/or the terms of the Special Licence, it will be resolved initially between the Archive the licence holder and the 'responsible' organisation.
19, 20 & 21 Agreement and Approval of the Special Licence	Please see the 'General Notes' above

***Further information**

Responsibility for the Licence holder's use of the data

Commissioning / funding	**Organisation's Representative**

organisation	
GSS	Head of Profession
Government Department	Head of Directorate or Division responsible for statistical analysis and research
University	Either: (i) Chair of the University Ethics Committee (ii) Director of Research (iii) Head of Department
Local Authority / Other bodies	Either: (i) Head of Directorate or Division responsible to the organisation for statistical analysis and research (ii) Person with authority to enter the organisation into a contract and with institutional responsibility for the actions of licence holder (iv) Person with the authority to take ultimate responsibility for the use of the data, the actions of the licence holder, breach of the terms of the Licence and any sanctions arising therefrom, i.e. the person who signs the Special Licence in this capacity will have the responsibility to enter their institution into an agreement that carries penalties for misuse and breach of the terms of the Licence that will impact both upon the institution and the licence holder.

APPENDIX F

Application for Access to Genotype Data and/or Biological samples from the Biomedical assessment of the 1958 birth cohort.

This form should be used to apply for samples, or to make requests that require links between samples and/or phenotypes and/or genotypes.

1. Details of all applicants

Principal Applicant Details	
Principal Applicant Name	
Position Held	
Affiliation	
Address line 1	
Address line 2	
Email	
Telephone	

Contact Person Details	
Name of Contact Person	
Position Held	
Affiliation	
Address line 1	
Address line 2	
Email	
Telephone	

Co-applicant Details	
Co-applicant 1 Name	
Co-applicant 1 Affiliation	
Co-applicant 2 Name	
Co-applicant 2 Affiliation	
Co-applicant 3 Name	
Co-applicant 3 Affiliation	
Co-applicant 4 Name	
Co-applicant 4 Affiliation	
Co-applicant 5 Name	
Co-applicant 5 Affiliation	

2. Project Details

Title of the project [less than 30 words]	
Proposed Project start Date	DD/MM/YYYY: //
Project finish Date	DD/MM/YYYY: //
Key words for application	

3. Brief description of project	
[10 key references maximum]	
1-2 sides of A4 Guidance	
Note	*Please ensure you have read the Guidance notes relating to the required description of the project.*

Continued: (Brief description of your project) -

Please include a lay summary of your project (150 words max. Guidance)

4. Funding Details	
Has the project been peer reviewed?	YES / NO
When was the project reviewed	DD/MM/YYYY: //
Has the project been funded?	YES / NO
Name of funding organisation	
Final Decision of the funders	
Funding start date	DD/MM/YYYY: //
Funding end date	DD/MM/YYYY: //

5. DNA	
Notes	*Your requests should be consistent with the project description in Section 3.*
Do you require DNA samples?	YES / NO If no please go to question 6, If YES please complete the following:

5.a Requested Sample Details	
Notes	*Your request should be consistent with the project description in Section 3.*
Will the project analyse samples from all available samples/subjects in the cohort study?	YES/NO
IF NO, please define the subset required?	
Quantity of DNA required	(µg) per sample/subject [Please note standard aliquots are 1µg at a concentration of 50ng/µl but larger quantities and concentrations are available on request.]
Minimum concentration required	(ng/µl) [Please note standard aliquots are 1µg at a concentration of 50ng/µl but larger quantities and concentrations are available on request.]
Number of subjects	
Is your request is for greater than 1 µg per sample/subject?	YES / NO
If yes please justify the size of the sample you have requested	

5.b Indicate whether the DNA is for	
Notes	*Your intentions should be consistent with the project description in Section 3.*
SNP analysis?	YES / NO
If so, how many SNPs?	Approx:
Micro-satellite analysis?	YES / NO
If so, how many microsatellites?	Approx:
Sequencing?	YES / NO
If so, what length?	Approx:
Structural DNA work (including copy number variation)?	YES / NO
Other?	Please specify:

5.c Please provide details of where DNA will be analysed	
Name of person responsible for analysis	
Laboratory address	

5.d. DNA preparation, storage and transport		
Are you happy to receive cell-line DNA?		YES / NO
If NO, please provide an explanation for your requirements		
Please provide a copy of the protocol(s) to be used for laboratory processing and analysis, including Q.A./Q.C. documentation		<These may be attached as a separate sheet to the application> Have you attached a copy of the protocol(s) to be used for laboratory processing and analysis? YES / NO
Notes		*If you are seeking access to a finite resource, protocols may be sent in confidence to external scientific peer-review.*
Are you aware that in order to obtain the DNA requested in this application, you are required to agree to return genotypes to enhance the 1958 BC resource?		YES / NO If yes, please sign and date below to confirm this agreement: Signature:_____ Date: _____
Please indicate if you require an embargo period before other users can access the data (up to 1 year from the date on which the data/samples are awarded)		YES/NO

6. Other Biological Samples		
Will the project require access to biological samples other than DNA?		YES / NO If no please go to question 7, If YES please complete the following:
Will the project process all available samples from the cohort study?		YES / NO IF NO, please define the subset required (this information is important, please read guidance):
Do you require Plasma?		YES / NO If yes, give the quantity of plasma required from each subject? If yes, please specify the preferred anticoagulant (EDTA, CPDA and citrate available)?
Do you require Serum?		YES / NO If yes, give the minimum quantity of Serum required from each subject?
Do you require Saliva?		YES / NO If yes, give the minimum quantity of Saliva required from each subject?
Do you require lymphoblastoid cell lines?		YES / NO

Continued: (6. Other Biological Samples) -		
Please specify what will be analysed		
What is the justification for the volume/quantity requested?		
Notes		*Your intentions should be consistent with the project description in Section 3.*
Please confirm you have looked at processing and storage history of serum, plasma and saliva samples and confirm they are suitable for your analysis		YES / NO
Notes		*Information about the storage history is available on the website (http://www2.le.ac.uk/projects/birthcohort/1958bc/open-calls/biosamples) or email Cohort1958@le.ac.uk for further information*
Please provide a copy of the protocol(s) to be used for laboratory processing and analysis		<These may be attached as a separate sheet to the application> Have you attached a copy of the protocol(s) to be used for laboratory processing and analysis? YES / NO
Notes		*If you are seeking access to a finite resource, protocols may be sent in confidence for external scientific peer-review.*
Are you aware that in order to obtain the samples requested in this application, you are required to return the results of assays generated under your project to enhance the 1958 BC resource		YES / NO If yes, please sign and date below to confirm this agreement: Signature:_____ Date: _____
Please indicate if you require an embargo period before other users can access the data (up to 1 year from the date on which the data/samples are awarded)		YES/NO

7. Genome wide genotype data		
Notes		*The genome wide genotype data from the 1958BC are held at the European Genome-phenome Archive (EGA) run by the European Bioinformatics Institute at the Wellcome Trust Genome Campus, Hinxton, UK. By default these genotype data come labelled with a binary indicator of sex, and a 12-level indicator of region of residence in Great Britain, so these do not need to be requested as extras.*
7.a: Does your project require access to genotype data from the available genome wide scans?		YES / NO
7.b: Does your project require access solely to genotype data available at the EGA?		YES / NO If YES: You should redirect your application the Consortium Data Access Committee of WTCCC (cdac@wellcome.ac.uk) If NO (that is you want anything else other than just the individual level genotypes with indicators of sex and region of residence), you should continue completing this form as the application will be dealt with by the Access Committee for CLS Cohorts (ACCC).
Notes		*It is only logical to respond to the following question (7c) if you have answered YES to question 7a and NO to question 7b.*

7.c: Which subsets of 1958BC data do you require from the GWA archive at EGA?		*This information is important, please read the guidance document.*

8. Other genotype data:		
Do you want any other genotype data generated from the 1958BC?		YES / NO
If YES, please specify the data you require.		
Were the genotypes you are requesting generated by a previous primary user of 1958BC samples or data?		YES / NO If yes, please indicate where those data are stored and under whose administration:
Have you discussed your proposal with that primary user?		YES / NO If yes, please outline their response:
Please specify which genotypes you need		

9. Other data:		
Notes		*Non-genotype data are held at the ESRC Data Archive at Essex University. By default, all samples and genotype data are released with indicators of sex and a twelve–level region of residence in Great Britain.*
Are these the only two variables you require?		YES / NO If yes, you do NOT need to complete this section, please proceed directly to question 10.
Notes		*Although data are an infinite resource, they may be sensitive and are potentially disclosive. Careful attention is therefore paid to ensuring that data are not released in unjustifiably large amounts. Any requests for specific variables must be carefully justified. This is particularly important if you request sensitive variables, or if you are proposing to undertake an analysis involving linkage of non-genetic (phenotype) data to genotype data or to DNA samples.*
If possible please detail the variables required:		
Notes		*This information is important, please read the guidance material.* *The data dictionary is available at: http://www.cls.ioe.ac.uk/datadictionary.*
Do you have any further enquiries regarding other data?		YES / NO IF yes, please contact Jon Johnson: j.johnson@ioe.ac.uk
Does your project require access to non-genotypic data other than sex and region of residence?		YES / NO If YES, *please indicate which variables and carefully justify your request*
Notes		*Your intentions should be consistent with the project description in Section 3.*

10. New variables created by you		
Will any new variables be derived or produced by you or your project team as a result of any data obtained from the 1958BC?		YES / NO / unsure If no, you do not need to complete this section. Please proceed directly to question 11.
Notes		*This information is important, please read the guidance material.*
If YES, please describe what variables you expect will be generated by your study		
If yes, are you aware that in order to obtain this data requested in this application, you should be willing to agree to return these variables to enhance the 1958BC resource?		YES / NO If yes, please sign and date below to confirm this agreement: Signature:_____ Date: _____
Please indicate if you require an embargo period before other users can access the data (up to 1 year from the date on which the data/samples are awarded to you)		YES / NO

11. Statistical analysis		
Who is responsible for statistical analysis of your project? Name		
Notes		*Please provide a single name of the person responsible for statistical analysis*
Affiliation		
Address line 1		
Address line 2		
Email		
Telephone		

Additional Info: Sample Size Calculations
Please provide a brief overview of sample size calculations for your proposed study:

11. Agreement

To be completed by the Principal Applicant:	
Can you confirm that you have read the above application?	YES / NO
Is the information contained in it is true to the best of your knowledge?	YES / NO
Do you understand that data and samples from the 1958BC resource cannot be used for commercial purposes?	YES / NO
Are you aware that if you, a member of your group, or your institution were to use these data for such a purpose *without* obtaining prior approval from the Access Committee for CLS Cohorts (AC3), you will be in breach of the material and/or data transfer agreements, and that this might result in you being excluded from using the 1958BC resource in the future?	YES / NO
Do you understand that if you undertake work that might potentially be viewed as commercial, it is your responsibility to seek the advice of the AC3?	YES / NO
Do you understand that you must not pass on any data or samples awarded, or any derived variables or genotypes generated by this application to a third party (i.e. to anybody that is not included in this list of applicants on this project, nor is a direct employee of one of these applicants)? (This would include any sharing of individual level data with a publically accessible archive).	YES / NO
Are you aware that any third party seeking to use data, samples, or derived variables or genotypes arising from this application must approach the AC3 to obtain access permission of their own?	YES / NO
Do you understand that if a problem arises involving any misuse of the 1958BC data or samples provided for this project - that violates any of the terms and conditions specified by the MTA or DTA that you have signed (as the principal applicant) will mean that you will be held responsible, and that this might result in you being excluded from using the 1958BC resource in the future?	YES / NO
Signature:_____ Date: _____ Print Name: _____	

Please send completed forms to:

Mrs Janet Jones
Access Committee for CLS Cohorts Secretariat
Department of Health Sciences
University of Leicester
Princess Road West
University Road
Leicester
LE1 7RH
UK
Tel : +44 (0)116 229 7232
Fax: +44 (0)116 229 7250
email: jrj3@le.ac.uk

Additional information for guidance:

Please note that assuming your application to the 1958 Birth Cohort resource is successful the final approval will be subject to the following stipulations:

1. Data and samples from the 1958BC resource cannot be used for commercial purposes and any commercial involvement would breach the basis on which the access has been awarded.

2. Third party sharing of either data or biosamples is strictly prohibited. Any third party seeking to use the data, samples or derived variables or genotypes must apply directly to the Access Committee for CLS Cohorts to obtain access permission in their own right.

3. The Access Committee requires that, where possible, individual level data items created de novo are made available to other users in accordance with contemporary best practice and taking appropriate account of ethico-legal restrictions and recognising any potential risks of disclosures of summary level genotypes[i]. If you believe that there is some reason that you can't meet this stipulation, please contact the Secretariat for the Access Committee.

4. For applications involving linked phenotype and genotype data it is important to note that once an award has been made, any future additions to the dataset (for example, if an additional linked phenotype variable is required) will have to be processed by the 1958 Birth Cohort Access Committee (Technical Review Team) and must comply with the original application. If you do need additional variables to be added, you should therefore inform the Secretariat of the Access Committee.

5. Applicants are reminded that the Terms and Conditions for the cohort explicitly forbid any attempt to identify individuals or to compromise or otherwise infringe the confidentiality of information on data subjects and their right to privacy.

6. **Incidental findings of clinical significance and potential benefit**

In signing their original consent forms for inclusion in the 1958BC Biomedical Survey (2002-2003), consenting participants agreed that they would not receive feedback about any individual genetic results: *"...no information found in the DNA will be given to me"* (NCDS Medical Follow-Up, Consent Form 2 – blood samples). In keeping with this wording the current policy of the ACCC is that *no* genotypic information (regardless of its nature) will be returned to cohort members.

To date, most informed commentators have seen this position as 'good practice' because nobody has really known how to interpret the clinical relevance of the genetic variants that have been identified: their effects have typically been rather small and there has been no agreed way in which to respond to the limited increases in risk they may convey. But in common with many of the world's major cohort studies and biobanks, the 1958BC recognises that national and international views of what constitutes 'best practice' might be about to change. For example, as outlined by a senior international commentator in the field[2], it is possible that in the future it may become mandatory to report genetic results to participants if they satisfy three key requirements:

(i) scientific validity (the genotyping is of adequate quality);
(ii) clinical significance (the disease or condition caused by the genetic variant is potentially serious) , and
(iii) potential benefit (*i.e.* a valid approach exists to prevent or cure the condition/disease of concern and that early knowledge of the genetic risk to which an individual is exposed could enhance the efficacy of that prevention/cure).

At present a change in what is seen as best practice remains no more than a hypothetical possibility, but findings that satisfy the three stated criteria are likely to become more common as the global scientific focus moves to full sequencing of genes and/or longer segments of DNA. The ACCC therefore wishes to help contribute to the national and international evidence-base on which any future strategic decisions might be made regarding policy for feeding back genetic results.

For this reason, **the ACCC now *requires* that if in the course of any analysis of DNA from any participant in the 1958BC, a genetic variant is found that could potentially be viewed as meeting all three of the criteria stated above, that information must be transmitted to the ACCC.**

At this stage this is no more than an exercise in collection of key data to assist us in developing an appropriate future strategy for the 1958BC – transmission of any information in this manner does not absolve the research group which generates the relevant finding from having their own internal policy to deal with this globally recognised problem. It is also important to ensure that your research group policy is consistent with the facts that: (1) at present NO genetic information can be returned to 1958BC participants; and (2) even if that policy were to change, all such contacts with cohort members would necessarily be undertaken by the Centre for Longitudinal Studies (contactable via ACCC). These requirements are immutable under any circumstances – even at the direction of an ethics committee that has reviewed your (the research group's) project.

---

1. (Policy for Use and Oversight of Samples and Data arising from the 1958 Cohort at http://www2.le.ac.uk/projects/birthcohort/oversight-committee)
2. Knoppers BM, Joly Y, Simard J, Durocher F. The emergence of an ethical duty to disclose genetic research results: international perspectives. *Eur J Hum Genet* 2006;14(11):1170-8.

APPENDIX G

1958 Birth Cohort Biosample Strategy Guidelines May 2013

Naveed Sattar and Paul Welsh
University of Glasgow

Helen Colhoun
University of Dundee

Susan Ring
University of Bristol

Objective of the scientific strategy guidelines

The 1958 Birth Cohort is a unique and powerful longitudinal epidemiological study, with tissue samples stored in biobanks, which will allow further biomarker and epidemiological work. Available tissue includes saliva, plasma and serum samples which are described in detail in Appendix 1. DNA and lymphoblastoid cell lines are also available from cohort members but are not covered by this document. Use of the samples is covered by Research Tissue Bank Ethical Approval (09/H1010/12) and requests to use the material are assessed by the Access Committee for CLS Cohorts (ACCC) (see http://www2.le.ac.uk/projects/birthcohort) . The tissue samples are a finite resource and the ethical approval requires that requests to access the material is subject to peer review. The material was collected during the Biomedical Sweep in 2003 and is the remains of samples analysed at that time. There has been little interest in the samples until late 2012 when the ACCC started to receive requests to access the material.

The objective of this document is to facilitate access to the 1958 stored tissue samples so that they get the widest possible usage while ensuring that scientific rigour is applied in selecting proposals that will yield data which are i) reliable ii) epidemiologically or clinically informative iii) novel. As such, applications will be considered in light of the cohort design; successful proposal should maximise the epidemiological strengths of the cohort, whilst also recognising limitations of the biobank (in terms of blood draw protocols, processing, storage, and sample availability).

This document provides a framework for addressing and determining the scientific rationale for access issues for biomarker work. This document does not prescribe rigid criteria because it is impossible to predict the nature of access requests or long-term trends in scientific interest. This document, whilst not exhaustive, sets a framework for making relevant decisions, giving some relevant examples where appropriate. This document has been developed to reflect current best practice and will be reviewed regularly to ensure it remains in line with current guidelines. The strategy also needs to reflect current funder policy and the ACCC will consult/update funders if there are any proposed changes to the strategy.

1. Use of the samples should be specifically relevant to the 1958 study

Applications to use 1958 samples should clearly demonstrate that the proposed study will make use of longitudinal data and cannot be carried out in samples obtained from another source. All data generated from samples will be returned to the 1958 cohort and made available to other users. Samples will only be issued under the terms of a material transfer agreement which includes the statement:

"It is a condition of access to the samples that information obtained from the samples (including any derived data, for example, derived haplotypes or the results of bioassays) is submitted to the University of Bristol for inclusion in the central 1958BC database. All genotypes, and all bioassay results that are important enough to be used in a publication must be returned to the 1958BC database. "

Recipients will also be required to return or destroy any unused material at the end of the project as requested by the ACCC under the terms of the material transfer agreement.

2. Scientific strength of the proposal, and potential impact

Critically, one must always ask whether a particular biomarker to be measured will answer a relevant and **meaningful question**. Using longitudinal studies as a cross-sectional resource is rarely impactful (aside from Mendelian Randomisation studies). Further, use of longitudinal data to investigate associations (hazard ratios, or risk ratios) must be justified on the grounds of potential clinical (or social) relevance. Which questions are generally meaningful in biomarker studies?

i) Clinical questions which might change the guidelines for clinicians, or give a clear public health message. Examples could include:
   a. **Disease diagnosis** e.g. HbA1c for diabetes or LFTs for NAFLD
   b. Vitamin D status in pregnant women and BMD in their children

ii) Clinical or social questions which might **risk stratify** patients e.g.
   a. Does NT-proBNP add informative to existing CVD risk scores?
   b. Do novel biomarkers improve prediction of clinical or social outcomes beyond established predictors

iii) **Disease pathogenesis**. Observational studies tend to be poor in investigating causality, even where impressive multivariable adjustment models are built. Wherever possible, proposals of this nature should consider whether a robust approach to causal identification can be applied, for example including whether the DNA resource can be combined with the proposal to use a Mendelian randomisation approach (assuming valid genetic instrumental variables are known and measured):

a. Do natriuretic peptides protect against diabetes?

iv) **Stratifying patients for therapy** based on phenotypes. Does a particular biomarker predict better or worse response to particular therapies?

3. Novelty of the scientific aims

Often the proposals with the most obvious and immediate scientific rigour will be the least novel studies; several cohorts may have conducted similar studies before. As such the balance between a proposal's strength (in terms of potential impact) and its novelty (which studies have measured the biomarker and related measures to outcomes before) is a key factor. If a proposal to measure a novel biomarker with little previous literature is interesting and potentially impactful, this must be considered in light of what is known regarding the biomarker (points below). Often, if a biomarker is particularly novel, a small pilot study may be useful prior to committing samples from the bioresource.

4. Biomarker characteristics; pre-analytical variables

Given the scarce nature of the bioresource, pre-analytical considerations as to whether a biomarker can be measured to give reliable results in the 1958 tissue samples are a key consideration (specific details for each sample type are provided in Appendix 1):

i) Sample processing: The 1958 blood samples were sent by post. The time spent with serum/plasma in contact with cells will have a significant impact on some biomarkers, but not others. Platelets release inflammatory factors, cells metabolise others, and the time spent at room temperature may adversely affect labile proteins. As such it should be noted that UK biobank have investigated pre-analytical characteristics of several of the more common biomarkers:

a. Glucose requires fast separation and assay to be conducted on first thaw.

b. **C-reactive protein** (**CRP**) is extremely robust to pre-analytical variables.

c. Limited existing data suggest metabolomics analysis may not be appropriate in samples not rapidly separated or at least within 24 hours

Given this, proposals must make it clear, with robust data to support the proposal, that the biomarkers to be measured will be reliably measured using the 1958 samples. This could be demonstrated with a pilot study, or published data, showing that sample processing time has no impact on the biomarker, or at least has a highly predictable effect (Passing-Blok regression, Bland-Altman plots etc). Pilots are always helpful before committing considerable time and money on novel biomarkers

ii) Freeze-thaw: The EDTA samples have not been previously thawed, whereas the citrate has. Many immunoassays, which measure based on antigenic structure rather than protein activity, are very robust to freeze thaw. This is likely to be the case for most biomarkers that are relatively unaffected by the sample processing time. Nonetheless, in supporting a proposal, data on the impact of freeze-thaw on a biomarker would be useful. In order to maximise use of the resource, it should be considered whether a previously thawed aliquot would be more appropriate to use (where possible) for a biomarker known to be robust to freeze-thaw.

iii) Sample type: There is more EDTA available than serum or citrate. The remaining serum aliquot is therefore important. Therefore, biomarkers which can be measured on EDTA should be in order to save the scarce serum resource for outstanding proposals. Very few non-haematological biomarkers are routinely measured in citrated plasma samples.

iv) Sample stability: All blood samples are stored at ≤-70^{o}C, so this issue is of limited relevance for biomarkers in the 1958 study.

5. Assay test platform

Assays should, where possible, be carried out using gold standard automated methods. In order of preference;

i) On an automated clinical chemistry/immunoassay platform in an accredited NHS laboratory, or a lab that participates in external quality assurance schemes for that assay

ii) On an automated platform in a laboratory using manufacturer recommended or internal quality control material

iii) Using single-plex assays such as ELISAs

iv) Using multiplex immunoassays

This list is intended as broad guidance, and there will be other potential assay methodologies. The gold standard for measuring vitamin D (25OHD2 and D3) is liquid chromatography tandem mass spectroscopy. Many aspects of this assay can be automated and carried out in NHS labs.

There is a broad trend towards use of multiplex assays to make optimal use of bioresources in epidemiology. Our own experience suggests that this technology should be used with caution. We have experience with Luminex (magnetic beads), Randox (bio-chips) and MSD (Multi-spot ELISA with electrochemiluminscence reporter) platforms. We have found:

i) Extra information comes at the cost of vastly reduced sensitivity and precision (higher CVs).

ii) Luminex beads system is rather sub-optimal for human blood samples; the beads tend to clog together making the assay method difficult/impossible to carry out within manufacturer recommended tolerances.

iii) The assay panel in multiplex assays are often of limited incremental value. Assaying C Reactive Protein (CRP) and Interleukin 6 (IL-6) in a study may be useful, but the incremental value of a dozen other cytokines may be limited or lack cost benefit, particularly when a majority are below the limit of sensitivity, or have limited or uncertain biological relevance. NB: multiplex assays often lead to reduced sensitivity for some tests and tend to lower CVs. Furthermore, where assay perform better e.g. MSD platform, there may be issues with respect to external

iv) Comparisons of data since some assays give results which are not externally comparable to values obtained by gold-standard methodologies, thereby required a conversion or "fiddle" factor.

Given the above, any proposal should be able to demonstrate that the assay they propose is sensitive enough to detect a signal (<20% CV as absolute and more desirable <10%) in a majority of the samples (commensurate with the aims). Ideally the platform/manufacturer used should be established in the literature to maximise the potential impact of the results, and minimise potential referee criticisms.

6. Assay test characteristics

This is a practical consideration, once a strong scientific case for a biomarker has been made in a proposal. An automated assay will have a dead volume (often ~200uL). For all assays the volume of sample consumed by the assay should also be considered in light of the potential impact of the study. If an EDTA sample has been previously thawed, the repeated use of this sample for other assays should be considered. If the volume remaining is too small for an automated assay it may remain sufficient for use in an ELISA assay by manual pipetting by a technician.

Often, multiple tests can be run on the same sample in automated platforms thereby maximising efficiency.

7. Global Discovery Versus Specific Hypothesis

All the above refers to specific tests of hypotheses; an alternative approach would be to reserve part of the resource for a more global discovery approach; specifically, it would be of interest across a wide range of disease states and phenotypes to acquire as much data as possible on the lipidome,

proteome and metabolome from high dimensional methods. For consideration might be mass spectroscopy (often semi-quantitative) and Nuclear magnetic resonance (NMR) based methods for quantitation of many small molecular weight metabolites and some peptides and proteins. Most experts in the field suggest if sufficient volume is available, the best approach for metabolomics is a combination of mass spectroscopy and NMR. Also for consideration are antibody-based arrays for high dimensional protein quantitation. Other methods to consider include proximal ligation assays for proteins, NMR based methods and mass spectroscopy methods for molecular species lipid analyses etc. Also one might consider serum micro RNAs worth detecting and quantifying.

Many of these approaches require relatively little volume (e.g. at least 600 serum metabolites can be detected and quantified with 120 ul, whereas other Mass spectroscopy platforms can yield potentially more than 1000 metabolites on 20 ul serum). However what is also true is that for many of the available platforms there is a surprising dearth of good data on the within person repeatability over short periods of time , the test re-test repeatability, pre-analytic effects on sensitivity and specificity and so on i.e. basic QC. For protein arrays etc, sensitivities may be particularly important to check since for some specific measurements high sensitivity single-plex ELISAs are employed (e.g. IL-6 in cohort studies) since conventional assays (and potentially arrays) cannot reliably pick up such low levels. Therefore before committing such a precious resource to any of these platforms careful consideration and possibly some pilot studies with less valuable samples are to be recommended. Furthermore, for some of these techniques, the statistical analyses can be very complex and in some cases, the best bioinformatics approach to analyse data, in particular data generated from mass spectroscopy, remains unclear. Finally, in all cases, whilst new techniques allow discovery science, the linkage of any measurements to pre-defined outcomes or to answer specific questions on disease pathology will help focus analyses.

SUMMARY

The 1958 tissue samples are a valuable resource but there are limitations regarding their suitability for some assays due to the sample processing history. Recommendations for ACCC for approving use of the samples are:

- Scientific strength of the proposal must justify use of 1958 cohort samples.
- Evidence must be provided to show methodology is appropriate given the processing history of the samples. Eg. Evidence from published literature or pilot data generated on samples processed in a similar manner.
- The assay test platform should have proven quality assurance measures in place.
- The methodology should include measures to ensure the quality of any remaining sample is not jeopardised and can be used in further assays which can be used on freeze thawed samples.
- At least one aliquot of each sample type should be reserved for future global discovery projects.

ACRONYMS

ACCC Access Committee for CLS Cohorts

CRP C-reactive protein

EDTA Ethylenediaminetetraacetic acid

ELISA Enzyme-linked immunosorbent assay

HbA1c glycosylated hemoglobin (*hemoglobin A1c)*

IL-6 Interleukin 6

MSD Meso Scale Discovery

NAFLD Non-alcoholic fatty liver disease

NMR Nuclear magnetic resonance

NT-proBNP N-terminal prohormone of brain natriuretic peptide

APPENDIX 1 - 1958 Birth Cohort Biological Sample Information

Table 1 – Summary of Samples collected and assays included in biomedical sweep from September 2002 to March 2004

Sample type	**Early Morning Saliva (Sarstedt salivettes)**	**Late morning Saliva (Sarstedt salivettes)**	**Citrated Plasma Residue**	**Plain serum residue**	**EDTA Plasma**	**CPDA Plasma (citrate-phosphate-dextrose-adenine)**
Maximum number of aliquots remaining	1 (varying volumes)	1 (varying volumes)	1 (varying volumes)	1 (varying volumes)	Up to 6 x 500 µl * + 1 varying volume	Up to 6 x 500 µl ** + 1 varying volume
Number of cases with at least 1 500ul sample remaining	6618	6618	7597	6400	8063	7848
Processing protocol	Transported by post at ambient temp. Frozen - at -80^{o}C in temporary storage, Shipped at ambient temperature to Germany for analysis. Refrozen on arrival. *No information regarding how samples were shipped back but currently stored at at -80^{o}C.*	Transported by post at ambient temp. Frozen - at -80^{o}C in temporary storage, Shipped at ambient temperature to Germany for analysis. Refrozen on arrival. *No information regarding how samples were shipped back but currently stored at at -80^{o}C.*	Shipped by post at ambient temp. 0.5ml of whole blood removed for analysis of glycosylated haemoglobin. Remainder centrifuged, aliquots frozen at -70^{o}C, transported frozen to Glasgow Royal Infirmary for analysis. Residue retained at -80^{o}C	Shipped by post at ambient temp. Centrifuged and the supernatant serum used for analysis in Newcastle. Residue retained at -80^{o}C	Shipped by post at ambient temp. Centrifuged and supernatant plasma stored in 0.5ml individually barcoded aliquots at -80^{o}C. Cell residues frozen and transported frozen to Bristol for DNA extraction.	Specific blood tube for production of lymphoblastoid cell lines. Shipped to Bristol by post at ambient temp. Centrifuged and plasma removed. Peripheral blood lymphocytes separated on a Ficoll gradient and cryopreserved for subsequent transformation into immortalised cell cultures. The supernatant

						plasma was sent to St George's Hospital Medical School (SGHMS) for aliquoting into 0.5ml individually barcoded tubes which were frozen at -80oC for long-term storage.
Processing Location	St George's Hospital Medical School then Germany	St George's Hospital Medical School then Germany	Royal Victoria Infirmary, Newcastle	Royal Victoria Infirmary, Newcastle	St George's Hospital Medical School	ALSPAC, University of Bristol then St George's
Days from taking sample to arrival in lab			1 day 18.9% 2 days 47.1% 3 days 24.2% 4 days 7.2% 5 days 1.5% >5 days 1.0%	1 day 18.9% 2 days 47.1% 3 days 24.2% 4 days 7.2% 5 days 1.5% >5 days 1.0%		Time to reach ALSPAC 1 day 17.8% 2 days 45.6% 3 days 24.9% 4 days 7.8% 5 days 1.9% >5 days 2.0%
Existing assays	Cortisol	Cortisol	glycosylated haemoglobin fibrinogen, tissue plasminogen activator, von Willebrand factor, C-reactive protein.	triglycerides, total and HDL cholesterol, total and allergen-specific immunoglobulin E, insulin-like growth factor 1	DNA	Lymphoblastoid cell lines
Current location	ALSPAC	UK Biobank	ALSPAC	UK Biobank	ALSPAC and UK Biobank	ALSPAC and UK Biobank

* See table 2 for more details

** See table 3 for more details

Table 2 – EDTA Plasma - Further details of number of 500 µl aliquots

Number of 500µl aliquots available	Number of cases
6	5110
5	2270
4	488
3	122
2	54
1	19

Table 3 – CPDA Plasma - Further details of number of 500 µl aliquots

number of 500µl aliquots available	Number of cases
6	7137
5	380
4	123
3	84
2	65
1	59

APPENDIX H

The Wellcome Trust Case Control Consortium Data Access Committee Membership

Prof Martin Bowbrow	University of Cambridge (*Chair*)
Prof Michael Parker	University of Oxford
Prof D Timothy Bishop	Leed Institute of Molecular Medi cine
Prof M Murtagh	University of Bristol

APPENDIX I

British 1958 birth cohort (1958BC)
Material Transfer Agreement for DNA or Biospecimens (version 6)

Reference number:

Start date:

Title of investigation:

Recipient Institution administrative contact
Name:
Address:
Tel:
Fax:
E-mail:

Principal Investigator at the Recipient Institution
Name:
Address:
Tel:
Fax:
E-mail:

Declaration

The Access Committee for CLS Cohorts (ACCC) has approved transfer to the Recipient Institution of Material from the 1958BC genetic or biospecimen resources at the ALSPAC Laboratory, University of Bristol, on the terms and conditions set out in this Collaborative Agreement. Accordingly, the Recipient Institution is willing to accept the Material and the University of Bristol is willing to accept data from the Recipient Institution in accordance with the provisions set out in this Collaborative Agreement.

The University of Bristol and the Recipient Institution hereby agree to be bound by the provisions set out in this Agreement.

Signed for and on behalf of the University of Bristol by its duly authorised representative

Signature: _____

Name: _____

Title: _____

Date: _____

Signed for and on behalf of the Recipient Institution

Signature: _____

Name: _____

Title: _____

Date: _____

Signature of Principal Investigator at the Recipient Institution

_____

Date: _____

Collaborative Agreement

Definitions

Providers of Material: The ALSPAC Laboratory, University of Bristol will provide the DNA or samples at the request of ACCC.
Material: includes the DNA or other biospecimens supplied and any derivatives or modifications thereof together with the associated descriptive data supplied, and the results of genotyping or other bioassays generated from individual or pooled samples of the DNA or biospecimens.
Publications: include but are not limited to articles published electronically or otherwise in peer-reviewed journals, reviews, books, posters and other written and verbal presentations of the research.
Research: the work being carried out by the Recipient Institution as detailed in Schedule 2.
Results: any data or information relating to the Materials which arises during the Recipient Institution's use of the Materials in the Research.

Terms and Conditions

A. In signing this Agreement, the Recipient Institution agrees and undertakes to:

1. That the Investigator and other relevant employees of the Recipient Institution involved in the Research have read and will abide by the "Policy for use and oversight of samples and data arising from the 1958 Birth Cohort (National Child Development Study)".

2. Use the Material in compliance with all applicable laws, governmental regulations and guidelines pertaining to research with the Material, including the Human Tissue Act 2004 and the MRC Guidelines on Human Tissue and Biological Samples for Use in Research (http://www.mrc.ac.uk/pdftissue_guide_fin.pdf).

3. Use the Material only for the purposes set out in the proposal attached at Schedule 2 and agreed by the ACCC.

4. Keep the Material under the immediate and direct control of the Principal Investigator.

5. Not transfer or make available the Material in whole or in part for any secondary distribution to any person other than those within the Principal Investigator's research group and for the declared and agreed use. (Should the Recipient Institution require another institution to assist them in their research, they must first obtain the consent in writing of the ACCC to such work being undertaken elsewhere, and then the third party must complete and sign a separate copy of this Material Transfer Agreement.)

6. Not use the Material or any parts thereof in or for the production of products for sale or for any commercial purpose.

7. Not hold the University of Bristol liable for any use by the Recipient Institution of the Material. The Recipient Institution agrees to indemnify and hold harmless the University of Bristol for any loss, claim, damage or liability of whatsoever kind or nature, which may arise from or in connection with this agreement, or the use, handling or storage of the Materials by the Recipient Institution.

8. Communicate promptly and in writing (E-mail is acceptable) to the University of Bristol any information regarding the quality of the Material or problems they may encounter with the Material or errors in the Material.

9. Not attempt to trace, contact or identify any individual member of the 1958 Birth Cohort or to recruit any cohort member to take part in any other survey.

10. Abide by the informed consent signed by cohort members during the 2002-2004 biomedical assessment as documented in Schedule 1.

11. Store the samples in a form which allows individual specimens to be removed in the event that a cohort member withdraws consent.

12. Take all reasonable steps to destroy the samples and products derived from the samples and data derived from the samples or derived products for any cohort member who withdraws consent. (Such requests from cohort members will be accepted only in writing and will normally be handled by the Centre for

Longitudinal Studies. The identifier for the specimen to be removed will be communicated to users by the University of Bristol and the Principal Investigator at the Recipient Institution will be asked to confirm in writing that the specimens and associated genetic data have been destroyed.)

13. The Recipient Institution will not use the Material in any experiments involving humans or animals and will not use the Material in contact with any cells or other materials to be infused into humans.

14. The 1958BC will own all Results directly relating to the Study Participants for the purposes of incorporation into the 1958BC resource. All other results generated by the Research shall be the property of the Recipient.

15. It is a condition of access to the samples that information obtained from the samples (including any derived data, for example, derived haplotypes or the results of bioassays) is submitted to the University of Bristol for inclusion in the central 1958BC database. All genotypes, and all bioassay results that are important enough to be used in a publication *must* be returned to the 1958BC database. The Recipient Institution will keep the ALSPAC laboratory, University of Bristol informed of the Results of the Research. The Recipient Institution will provide the ALSPAC laboratory, acting on behalf of the 1958BC, with an appropriately documented electronic copy of the Results before publication in any form or within 12 months of the completion of the Research whichever is the sooner. There will be accompanying documentation sufficient to identify the genotype (eg chromosomal location of the genetic variants) or bioassays tested, the interpretation of the coded results, and a brief description of the methods used. The format for this report will be agreed between the Recipient Institution and the ALSPAC laboratory, University of Bristol. Where necessary, the timing of lodgement, and the duration of any subsequent embargo on their use by others (maximum one year), can be agreed between the applicants and the ACCC. At the discretion of the ACCC the data may be lodged with the UK Data Archive. Applicants must supply adequate documentation concerning new variables (including statistical programs) to permit their use by others in future analyses of the data.

16. The Recipient Institution will acknowledge 1958BC and the funders and, where appropriate, will include as authors specific individuals identified by the ACCC who have played a substantial scientific role (as would be defined in a standard publication policy) in the generation of the Material used in a specific publication based on 1958BC data, samples or results. The secretariat to the ACCC must be informed of all research papers based wholly or partly upon the Material.

17. Inform the press offices at the Wellcome Trust and Medical Research Council prior to any media publicity.

18. Provide reports of progress or any other nature as requested by the ACCC, and notify the ACCC of any significant delays in completing the research proposed in schedule 2.

19. Return or destroy the Material at the end of the project as requested by the ACCC.

20. Recipient Institutions will be expected to meet all the costs of sample handling, specimen transport and data preparation in relation to their study.

B. This Agreement does not restrict the rights of the 1958 Birth Cohort, or those institutions authorised to act on its behalf, to distribute the Material to other institutions or to publish any document relating to this Material.

C. The Recipient Institution warrants that it has full legal authority in the country where the accompanying data will be processed to receive, store and process such data, to use it for the purpose(s) for which it has been collected, as set out in Schedule 1, and will comply with the Data Protection Act 1998.

D. Either party may terminate this Agreement for any reason on 30 days prior written notice to the other. Termination of this Agreement for any reason shall not relieve the Recipient Institution of its obligations under this Agreement.

E. This Agreement shall be governed by and construed and interpreted in accordance with the laws of England and the parties hereby submit to the non-exclusive jurisdiction of the Courts of England.

Schedule 1 – Consent

All protocols, information sheets and consent forms for the ongoing fieldwork were approved by the SouthEast MREC in August 2002 (ref: MREC 01/1/44). An information booklet was sent in advance and signed consent was obtained by the nurse at the time of blood sampling. Cohort members were asked to consent separately to venepuncture, storage of plasma, extraction of DNA and immortalisation of cell lines. The sections relevant to the blood collection and genetic resource are as follows:

CONSENT FORM 2 – Blood samples
I, *(name)* _____

a) Give my consent to _____ (qualified nurse) to collect a sample of my blood to be tested for cholesterol, glycosylated haemoglobin, fibrinogen, total and allergen-specific IgE. I understand that the blood samples and related information will be coded and used anonymously for non-commercial research purposes only, and will not be tested for HIV.
Signed _____ *Date* _____

b) Give my consent to storage of frozen portions of my blood sample for use in future medical research studies of the causes, diagnosis, treatment or outcome of disease. I understand that the blood samples and related information will be coded and used anonymously for non-commercial research purposes only, and will not be tested for HIV. I understand that I may withdraw this consent at any time by contacting the investigators in writing, without giving any reasons.
Signed _____ *Date* _____

c) Give my consent to extraction and storage of DNA from my blood sample for use in future medical research studies of the causes, diagnosis, treatment or outcome of disease. I understand that the DNA samples and related information will be coded and used anonymously for non-commercial research purposes only, and that no information found in the DNA will be given to me. I understand that I may withdraw this consent at any time by contacting the investigators in writing, without giving any reasons, and the DNA extracted from my blood samples will then be destroyed and any genetic data obtained from it will be deleted.
Signed _____ *Date* _____

d) Give my consent to storage of white blood cells for future creation of cell cultures. I understand that these cells will provide a renewable source of DNA for use in future medical research studies of the causes, diagnosis, treatment or outcome of disease. I understand that the cells, DNA samples and related information will be coded and used anonymously for non-commercial research purposes only, and that no information found in the DNA will be given to me. I understand that I may withdraw this consent at any time by contacting the investigators in writing, without giving any reasons, and the cell cultures and DNA obtained from them will then be destroyed.
Signed _____ *Date* _____

Schedule 2
[Copy of the application submitted to the Oversight Committee and any subsequent correspondence, including the letter of approval from the chairman, which may include specific additional conditions which the Recipient must follow in dealing with the Material.]

APPENDIX J

THE WELLCOME TRUST CASE-CONTROL CONSORTIUM

DATA ACCESS AGREEMENT

This agreement governs the terms on which access will be granted to the genotype data generated by the Wellcome Trust Case-Control Consortium.

In signing this agreement, You are agreeing to be bound by the terms and conditions of access set out in this agreement.

For the sake of clarity, the terms of access set out in this agreement apply both to the User and the User's Institution (as defined below). User Institution and User are referred to within the agreement as "You" and "Your" shall be construed accordingly.

Definitions:

Consortium means the Wellcome Trust Case-Control Consortium, a group of Wellcome Trust-funded investigators, a list of which can be found on the study website www.wtccc.org.uk.

Data means all and any human genetic data obtained from the Consortium.

Data Subject means a person, who has been informed of the purpose for which the Data is held and has given his/her informed consent thereto.

User means a researcher whose User Institution has previously completed this Data Access Agreement and has received acknowledgement of its acceptance.

Publications means, without limitation, articles published in print journals, electronic journals, reviews, books, posters and other written and verbal presentations of research.

User Institution means the organisation at which the User is employed, affiliated or enrolled.

Terms and Conditions:

In signing this Agreement:

1. You agree to use the Data only for the advancement of medical research, according to the consent obtained from sample donors.

2. You agree not to use the data from the 1958 British Birth Cohort or any part thereof for the creation of products for sale or for any commercial purpose.

3. You agree to preserve, at all times, the confidentiality of information and Data pertaining to Data Subjects. In particular, You undertake not to use, or attempt to use the Data to compromise or otherwise infringe the confidentiality of information on Data Subjects and their right to privacy.

4. You agree not to attempt to link the data provided under this agreement to other information or archive data available for the data sets provided, even if access to that data has been formally granted to you, or it is freely available without restriction, without specific permission being sought from the relevent access committees.

5. You agree not to transfer or disclose the Data, in whole or part, or any identifiable material derived from the Data, to others, except as necessary for data/safety monitoring or programme management. Should You wish to share the Data with a collaborator outwith the same Institution, the third party must make a separate application for access to the Data.

6. You agree to use the data for the approved purpose and project described in your application; use of the data for a new purpose or project will require a new application and approval.

7. You accept that Data will be reissued from time to time, with suitable versioning. If the reissue is at the request of sample donors and/or other ethical scrutiny, You will destroy earlier versions of the Data.

8. You agree to abide by the terms outlined in the Consortium 'Publications Policy' as set out in Schedule 1.

9. You agree to acknowledge in any work based in whole or part on the Data, the published paper from which the Data derives, the version of the data, and the role of the Consortium and the relevant primary collectors and their funders. Suitable wording is provided in the Publications Policy given in Schedule 1.

10. You accept that the Consortium, the original data creators, depositors or copyright holders, or the funders of the Data or any part of the Data supplied:

    a) bear no legal responsibility for the accuracy or comprehensiveness of the Data; and
    b) accept no liability for indirect, consequential, or incidental, damages or losses arising from use of the Data, or from the unavailability of, or break in access to, the Data for whatever reason.

11. You understand and acknowledge that the Data is protected by copyright and other intellectual property rights, and that duplication, except as reasonably required to carry out Your research with the Data, or sale of all or part of the Data on any media is not permitted.

12. You recognise that nothing in this agreement shall operate to transfer to the User Institution any intellectual property rights relating to the Data. The User Institution has the right to develop intellectual property based on comparisons with their own data.

13. You accept that this agreement will terminate immediately upon any breach of this agreement by You and You will be required to destroy any Data held.

14. You accept that it may be necessary for the Consortium or its appointed agent to alter the terms of this agreement from time to time in order to address new concerns. In this event, the Consortium or its appointed agent will contact You to inform You of any changes and You agree that Your continued use of the Data shall be dependent on the parties entering into a new version of the Agreement.

15. You agree that you will submit a report to the Consortium Data Access Committee, if requested, on completion of the agreed purpose. The Consortium Data Access Committee agrees to treat the report and all information, data, results, and conclusions contained within such report as confidential information belonging to the User Institution.

16. You accept that the Data is protected by and subject to international laws, including but not limited to the UK Data Protection Act 1998, and that You are responsible for ensuring compliance with any such applicable law. The Consortium Data Access Committee reserves the right to request and inspect data security and management documentation to ensure the adequacy of data protection measures in countries that have no national laws comparable to that which pertain in the EAA.

17. This agreement shall be construed, interpreted and governed by the laws of England and Wales and shall be subject to the non-exclusive jurisdiction of the English courts.

SCHEDULE 1

Publications Policy

The primary purpose of the Wellcome Trust Case-Control Consortium (WTCCC) is to accelerate efforts to identify genome sequence variants influencing major causes of human morbidity and mortality, through implementation and analysis of large-scale genome-wide association studies. Additional objectives include the development and validation of informatics and analytical solutions appropriate to the scale and nature of the project, as well as use of the data generated to answer important methodological and biological questions relevant to association studies in general, and in the UK in particular.

The Consortium anticipates that data generated from the project will be used by others, such as required for developing new analytical methods, in understanding patterns of polymorphism and in guiding selection of markers to map genes involved in specific diseases. A more detailed list of the WTCCC aims are provided on the WTCCC website – www.wtccc.org.uk

Authors who use data from the project must acknowledge the WTCCC using the following wording "*This study makes use of data generated by the Wellcome Trust Case-Control Consortium. A full list of the investigators who contributed to the generation of the data is available from www.wtccc.org.uk. Funding for the project was provided by the Wellcome Trust under award 076113 and 085475*" and cite the relevant primary WTCCC publication (details of which can be found on the WTCCC website).

Users should note that the Consortium bears no responsibility for the further analysis or interpretation of these data, over and above that published by the Consortium.

For and on behalf of User:

Name of Applicant(s): _____

_____

_____

_____

Signature of Applicant(s): _____

_____

_____

_____

Date: _____

For and on behalf of User Institution:

Signature of Institutional or Administrative Authority: _____

Print name: _____

User Institution: _____

Date: _____

WHEN SUBMITTING THIS DOCUMENT, PLEASE INCLUDE ALL PAGES OF THE AGREEMENT WITH THIS SIGNATURE PAGE

APPENDIX K

NCDS Medical follow-up

CONSENT BOOKLET – OFFICE COPY

National Centre for Social Research

Operations Department
100 Kings Road, Brentwood
Essex CM14 4LX
Telephone 01277 200 600
Fax 01277 214 117

APPENDIX K

P2107

NCDS Medical follow-up
CONSENT BOOKLET – OFFICE COPY

Please use capital letters and write in ink

NAME/ADDRESS – WRITE IN:

RESPONDENT NAME:
ADDRESS:

POSTCODE:

ATTACH SERIAL NUMBER BAR CODE LABEL:

1. Nurse number

2. Date schedule completed — DAY / MONTH / YEAR

3. Full name (of person tested) _____

Name by which GP knows person (if different) _____

4. Sex — Male 1 / Female 2

5. Date of birth: DAY / MONTH 0 3 / YEAR 1 9 5 8

6. GP NAME AND ADDRESS
Dr: ……
Practice Name: ……
Address: ……
……
Town: ……
County: ……
Postcode: ……
Telephone no: ……

7. NURSE USE ONLY

GP address complete	1
GP address incomplete	2
No GP	3

8. SUMMARY OF CONSENTS – RING CODE FOR EACH ITEM

	YES	NO
a) Vision tests to GP	01	02
b) Hearing tests to GP	03	04
c) Height and Weight to GP	05	06
d) Waist and Hip measurement to GP	07	08
e) Blood pressure to GP	09	10
f) Lung function to GP	11	12

	YES	NO
g) Sample of blood to be taken	13	14
h) Blood sample results to GP	15	16
i) Blood sample for storage	17	18
j) Blood sample for DNA extraction	19	20
k) Blood sample for cell cultures	21	22
l) Blood sample result to respondent	23	24

CONSENT FORM 1 - Measurements

I, (name) _____

give my consent to _____(qualified nurse)

to measure the following:

		Ring one code on each line	
a)	Tests of near and distant vision	Yes	No
b)	Blood pressure and pulse rate	Yes	No
c)	Pure tone audiometry tests of hearing threshold	Yes	No
d)	Standing and sitting height	Yes	No
e)	Body weight	Yes	No
f)	Waist and hip circumferences	Yes	No
g)	Lung function using a spirometer	Yes	No

I am willing to complete a structured interview about mental health Yes No

I have read the letter of introduction and the information leaflet about the development stage of the medical examination of the National Child Development Study. I have discussed any outstanding questions with the nurse named below and I wish to participate in the examination. I understand that I can stop the interview and examination at any point or decline any part of it, and that all information will be treated in the strictest confidence and used for research purposes only.

Signed _____ Date _____

Countersignature by nurse

I confirm that I have explained the nature of the proposed studies to the person named above and have left a copy of the information sheet and this consent form with them for future reference.

Signed _____(Nurse) Date _____

CONSENT FORM 2 – Blood samples

I, (name) _____

a) Give my consent to _____ (qualified nurse) to collect a sample of my blood to be tested for cholesterol, glycosylated haemoglobin, fibrinogen, total and allergen-specific IgE. I understand that the blood samples and related information will be coded and used anonymously for non-commercial research purposes only, and will not be tested for HIV.

Signed _____ Date _____

b) Give my consent to storage of frozen portions of my blood sample for use in future medical research studies of the causes, diagnosis, treatment or outcome of disease. I understand that the blood samples and related information will be coded and used anonymously for non-commercial research purposes only, and will not be tested for HIV. I understand that I may withdraw this consent at any time by contacting the investigators in writing, without giving any reasons.

Signed _____ Date _____

c) Give my consent to extraction and storage of DNA from my blood sample for use in future medical research studies of the causes, diagnosis, treatment or outcome of disease. I understand that the DNA samples and related information will be coded and used anonymously for non-commercial research purposes only, and that no information found in the DNA will be given to me. I understand that I may withdraw this consent at any time by contacting the investigators in writing, without giving any reasons, and the DNA extracted from my blood samples will then be destroyed and any genetic data obtained from it will be deleted.

Signed _____ Date _____

d) Give my consent to storage of white blood cells for future creation of cell cultures. I understand that these cells will provide a renewable source of DNA for use in future medical research studies of the causes, diagnosis, treatment or outcome of disease. I understand that the cells, DNA samples and related information will be coded and used anonymously for non-commercial research purposes only, and that no information found in the DNA will be given to me. I understand that I may withdraw this consent at any time by contacting the investigators in writing, without giving any reasons, and the cell cultures and DNA obtained from them will then be destroyed.

Signed _____ Date _____

CONSENT FORM 3

Saliva sample

I, (name) _____ give my consent to use of samples of my saliva for tests of cortisol and future medical research studies of the causes, diagnosis, treatment or outcome of disease. I understand that the saliva samples and related information will be coded and used anonymously for non-commercial research purposes only, and will not be tested for HIV. I understand that I may withdraw this consent at any time by contacting the investigators in writing, without giving any reasons.

Signed _____ Date _____

Consent to send results to GP

I, (name) _____ wish these results to be sent to my general practitioner so that they can be used to help monitor my health. I understand that my GP may wish to include the results in any future report about me:

		Ring one code on each line		
a)	Vision test results	Yes	No	Not measured
b)	Blood pressure and resting pulse rate	Yes	No	Not measured
c)	Hearing test results	Yes	No	Not measured
d)	Height, weight and measures of body size	Yes	No	Not measured
e)	Lung function test results	Yes	No	Not measured
f)	Blood test results for blood cholesterol and glycosylated haemoglobin	Yes	No	Not measured

Signed _____ Date _____

CONSENT FORM 4 – Archiving of data and Consent to obtain information from National Health Service medical records

I, (name) _____

a. Give my consent for the results of measurements and laboratory tests carried out on me as part of the medical examination of the National Child Development Study to be deposited at the Economic and Social Research Council Data Archive, as part of the National Child Development Study dataset. I understand that the archived information will be coded and used anonymously for research purposes only, and will not include my name or address.

Signed _____Date _____

b. Give my consent to use of information from my National Health Service medical records in future medical research studies of the causes, diagnosis, treatment or outcome of disease. I understand that the information obtained by the investigators will be coded and used anonymously for research purposes only, and will not include my name or address. I understand that I may withdraw this consent at any time by contacting the investigators in writing, without giving any reasons.

Signed _____Date _____

Appendix L

Returning Individual Genetic Research Findings in the Context of the National Child Development Study (1958 Birth Cohort): A Briefing Paper

Susan E. Wallace[1]

July 2012

First published in 2012 by the
Centre for Longitudinal Studies
Institute of Education, University of London
20 Bedford Way
London WC1H 0AL
www.cls.ioe.ac.uk

The Centre for Longitudinal Studies (CLS) is an ESRC Resource Centre based at the Institution of Education. It provides support and facilities for those using the three internationally-renowned birth cohort studies: the National Child Development Study (1958), the 1970 British Cohort Study and the Millennium Cohort Study (2000). CLS conducts research using the birth cohort study data, with a special interest in family life and parenting, family economics, youth life course transitions and basic skills.
The views expressed in this work are those of the author and do not necessarily reflect the views of the Economic and Social Research Council. All errors and omissions remain those of the author.

This document is available in alternative formats.
Please contact the Centre for Longitudinal Studies.
tel: +44 (0)20 7612 6875
email: info@cls.ioe.ac.uk

Acknowledgements

I would like to acknowledge the assistance of Dr Neil Walker in the preparation of this paper. I am also grateful to Dr Walker, ProfJane Elliott and Dr Madeleine Murtagh for their helpful comments on earlier drafts.

1. Introduction

The Wellcome Trust Case Control Consortium (WTCCC) is planning, together with industry partner Illumina, to develop a chip to genotype every possible SNP in the coding portion of the genome. The level of detail available on the chip will be similar to that of exome sequencing. This means that researchers using this chip may find information, whether as a direct result of their research or incidentally, that could have implications for participants. The data from studies using this exome chip may be available as early as June 2012. As this chip will use WTCCC data, and therefore data from the National Child Development Study (hereafter, the 1958 Birth Cohort (1958BC)), the potential impact of these finding on cohort members needs to be considered. In addition, genome-wide association study (GWAS) data is already available. In the WTCCC2 project, 3,000 1958BC participants were used as a national control resource for 2 GWAS chips. Results from studies using that data are also imminent. Therefore, it is timely for the Centre for Longitudinal Studies (CLS) to consider its position on this issue.

There are many different definitions of a 'result' or a 'finding' in the context of genomic research studies and population biobanks. Results could be data from the measurements taken when a person is joining a biobank or cohort study; many projects give back blood pressure readings or results from eye tests. If a dangerously high blood pressure reading, for example, is noted, usually the project team will advise the person to see their doctor or seek urgent medical treatment for them (Boddington *et al.*, 2011). The provision of this information has been seen as unproblematic in the main and for some as a recognised duty of care or rescue (Knoppers, 2009).

Concerns have arisen with results that could have implications for the health or well-being of study participants, Health-related Findings (HRFs). Such findings can be "…relevant to the research question (a 'pertinent' finding) or completely unrelated (an 'incidental' finding) (Opinion Leader, 2012). One subset of findings within this broad category of HRFs is the genetic variation information that is increasingly being discovered through the use of whole exome and whole genome sequencing methods. These individual genetic research (IGR) findings identified by secondary researchers through their use of 1958BC data are the focus of this paper.

Traditionally, large-scale population based cohort studies have not given IGR findings back to their participants (Wallace and Kent, 2011). However, with scientific knowledge progressing, it has been acknowledged that this is no longer a black and white issue. There are now detailed discussions as to whether or not participants in research should be given such information. From the individual's perspective, they will need to decide whether they will use findings to improve their health or lifestyle. Knowing their predisposition could influence their decisions about their future health, their reproductive choices and the choices of their children and other relatives. There are also some people who do not want to know, and this should also be respected. Table 1 shows some of the ethical and practical arguments for and against.

Table 1: Arguments for and against providing IGR findings to participants

Arguments for provision of IGRs	Arguments against provision of IGRs
Individuals have a 'right' to information about themselves;	Participants may not understand the information being given to them and it may make them unnecessarily anxious
There may be a positive health outcome for participants	False positives may result in individuals seeking interventions unnecessarily
It would increase trust in researchers	The findings might not be validated
It would increase individuals' desire to participate in research	There may not be an intervention available to help the individual
It would serve as reciprocity for participating in research	The process can be time consuming and expensive
It shows respect for participants	Research studies may not have the expertise necessary (e.g., genetic counsellors on staff)
	Research findings are not intended to clinically benefit individuals

There has been significant activity within the academic and funding communities to seek out evidence, from the scientific and lay communities, and use it to help craft policies that meet the needs of stakeholders. A recent study commissioned by the Wellcome Trust and the Medical Research Council has shown that the individuals sampled were overwhelmingly in favour of receiving HRFs when a condition is serious and treatable (Opinion Leader, 2012). This is echoed by other research regarding returning IGR findings (Bollinger *et al.*, 2012). Yet, many difficult questions remain, including whether, by which process, by whom, and which findings should be returning to research participants.

In the UK, a policy from The Wellcome Trust and the Medical Research Council on this issue is expected in autumn 2012. In the interim, this briefing paper outlines some points to consider including:

- What issues to focus on when deciding whether or not to return IGR findings from secondary researchers;
- Examples of current practices by other research studies;
- A suggested process by which decisions can be made as to which findings to return, and;
- Some examples of findings specific to the 1958BC and the factors that might be considered in deciding whether or not to return these particular findings and
- Examples of how other cohorts are dealing with IGR findings.

2. Points to consider when deciding whether or not to return results

While the pervading opinion reported by the academic community may be supportive of disclosing HRFs and IGR findings, those who are in charge of a study will need to decide whether or not it is appropriate in their context. The process for deciding involves several considerations.

2.2. Original consent language

A first step is to examine the original consent materials to determine to what the participants agreed, so as to decide what actions need to be taken. The 2002 NCDS Biomedical Survey was the first sweep in which genetic material was collected from the 1958BC. The Survey consent materials included information on the collection of 'medical' measurements as well as blood and saliva. Measurements included height, weight and other body measures, blood pressure, lung function, eyesight tests, and hearing tests. A wellbeing questionnaire was given if agreed. Saliva and a blood sample were also taken.

Table 2 shows the language used in the consent forms used for the Survey. It is clear from these excerpts that participants were informed that they would not receive any further information from analyses of the DNA. Measurement data would be fed back, if desired, and participants were given the option to choose which data they wished to see. The Survey also allowed for the possibility of involving a participant's general practitioner (GP); a written summary of the measurements could be sent to the GP, with the presumed aim of using that information in clinical care.

Reviewing this consent language shows two things. First, if the CLS decides not to return findings, this decision can be substantiated by the consent agreement. However, many groups are now reconsidering their policies on returning findings, even after having said no results would be returned. If the CLS decides it wishes to change its position, this may require re-consenting the cohort. Second, if IGR findings are to be fed back, the CLS has already 'built-in' the possibility of using GPs as part of the return process (see 2.5.2).

2.3. Re-consent

The CLS may need to investigate whether or not to re-consent the cohort. The first consideration is whether re-consent is required because the new use of the participant's data is significantly different from the original consented use. If not, then going through a re-consent process may be seen as harassing. As Steinsbekk et al point out, "[p]articipants should be asked again to re-consent only when there is a relevant ethical difference in the activities planned…" from those in the original consent form. (Steinsbekk and Solberg, 2011), p. 8. In the case of the 1958BC, this new use, the detection and return of IGR findings, can be argued as activities that are significantly outside the original consent, thus warranting re-consent.

Table 2. Language regarding feedback of information from 2002 NCDS Biomedical Survey consent materials

Consent Material	Text
Information Leaflet (participation in survey)	"…the nurse will give you a written summary of the measurements she has made on the day." "We will not send any information to your doctor (GP) without your permission, but if you agree, the results of all your measurements and laboratory tests will be sent to your doctor in a separate letter."
Information Leaflet (collection of genetic material/conduct of genetic studies)	*No applicable text included*
Consent Booklet (Office Copy)	"8. Summary of consents h) Blood sample results to GP [yes/no] l) Blood sample result to respondent [yes/no]"
Consent Form 1 – Measurements	*No applicable text included*
Consent Form 2 – Blood Samples	"c) Give my consent to extraction and storage of DNA from my blood sample…. …and that no information found in the DNA will be given to me." "d) Give my consent to storage of white blood cells for future creation of cell cultures. …and that no information found in the DNA will be given to me."
Consent Form 3 – Saliva Sample	*No applicable text included*
Consent Form 3 – Consent to send results to GP	"I, (name) ____wish these results to be sent to my general practitioner so that they can be used to help monitor my health. I understand that my GP may wish to include the results in any future report about me. a) Vision test results Yes No Not measured b) Blood pressure and resting pulse rate Yes No Not measured c) Hearing test results d) Height, weight and measures of body size Yes No Not measured e) Lung function test results Yes No Not measured f) Blood test results for blood cholesterol and glycosylated haemoglobin"
Consent Form 4 – Archiving of data and Consent to obtain information from National Health Service medical records	*No applicable text included*

There is some empirical evidence that shows that participants prefer to be re-consented rather than find that the study has gone ahead with a new use of their data without consultation. Ludman et al (Ludman *et al.*, 2010) explored participants' attitudes towards re-consent for a new use of their data, deposition in the NIH dbGap database. They gave participants three options: (1) seeking individuals' consent for the new use of data; (2) using an opt-out method, where participants are only asked to respond if they do not wish their data to be used in the new way; or (3) using a notification method where participants are notified that the data would be used in the new way. They found that participants overwhelmingly approved of being asked for re-consent and were very disapproving of either the opt-out or notification methods. This seems to be the prevailing attitude; people feel that by asking for their consent, researchers are acting transparently, as well as showing respect for that participant's contribution to the research. (Trinidad *et al.*, 2011)

However, deposition of data into a database could be seen as potentially risky by participants, which may be why they wanted to be asked rather than simply told. Participants might feel differently about knowing they could potentially receive beneficial health data. As this can been seen as a positive change in policy, rather than a negative one, the CLS might be able to consider an opt-out or notification only method.

2.4. Costs

Budgetary issues need to be considered. First, there is the estimated cost of re-consent procedures. These costs will differ depending on the procedures chosen but can be expensive and extra funding may be needed. There is precedence for cost determining whether genetic research will be done; Clayton and McGuire report that "the Centers for Disease Control and Prevention decided not to proceed with genetic research in the National Health and Nutrition Examination Survey III in part because the cost of obtaining adequate consent was estimated to be in the millions of [US] dollars" (Clayton and McGuire, 2012).

The estimated cost of actually giving back results also needs to be considered. Procedures will vary and evidence is limited, but Christensen et al report that the cost of giving CDKN2A research results to 19 melanoma survivors averaged over US$1,300 per disclosure, including personnel time, genotype confirmation, educational materials and counselling (Christensen *et al.*, 2011). This example was a one-off process; one can assume that the CLS will be expecting IGR findings over the long term. The CLS will need to consider what infrastructure it will use and how the funding for this would be sustained. A decision will need to be made as to who will bear these costs – the cohort, the secondary researcher or both.

2.5. Potential loss of cohort members

If re-consent is carried out, there is the real possibility that a number of cohort members will decide not to agree to the new use of their data and therefore their data will not be eligible for genetic studies. In addition, a small number of cohort members may decide to withdraw from the study altogether. The CLS will need to decide what the impact of this might be on genetic studies, as well as on the study in general. An important issue that should not be over looked when discussing re-consent is that although a sample of 8406 consented to DNA extraction from their blood sample in 2002/3, of these *only* 7407 completed the age 50 sweep of NCDS. This is because 36 had died since the biomedical, 48 had emigrated, 321 were contacted and refused, and 595 could not be traced. The best estimate of the target sample for re-consenting would therefore be 7407. Clearly CLS could aim to re-consent

some cohort members who had not taken part in the age 50 sweep of NCDS in 2008, however further cohort members will have died and emigrated since 2008.

As noted earlier, currently the samples are not consented for commercial or non-medical use. Table 3 shows how many samples would be available if the cohort were re-consented for these different areas of research. Four hypothetical scenarios are presented, based on different response rates and agreement rates. What is key here is that even with relatively high response rates and agreement rates, following a re-consent exercise, adopting an 'opt-in' procedure, the sample available for commercial and non-medical research is likely to fall below 6000. In addition if we focus on the third row, which presents the numbers responding under the different scenarios, we have an estimate of those who would respond to the question about whether they wanted IGRs returned to them.

Table 3. Impact on the numbers of samples available for genetic and other research based on estimated re-consent figures

	Scenario A 85% RR* 95% Agree±	**Scenario B 85% RR 90% Agree**	**Scenario C 80% RR 95% Agree**	**Scenario D 70%RR 90% Agree**
Samples available for medical and non-commercial research (current position)	8406	8406	8406	8406
Number of cases with consent who took part in the age 50 survey in 2008 (i.e. target sample for reconsent process)	7407	7407	7407	7407
Number of cases responding to request for re-consent	6296	6296	5926	5926
New sample re-consented for commercial and non-medical	5981	5666	5629	4666
Number withdrawing consent	252	252	237	207

* RR = Response Rate
±Agree = Agreement to commercial and non-medical research

Based on these assumptions, a considerable number of cohort members are unlikely to respond to a re-consent process, given that this would be taking place a decade after the original consent was obtained. In addition, as participants and their samples will be lost to the *whole study* each time re-consent is taken, the CLS will need to decide if, and for what, it wants to re-consent its participants.

2.6. Infrastructure/processes

There are infrastructure or process questions that need to be considered. In other words, the CLS will need to decide first, how it is going to decide which findings, and second, how it is going to go about returning those findings. There are limited examples of how these are being done by other studies at this time.

2.6.1. Infrastructure/process for deciding which findings to return

There are several ways of approaching the decision of which findings to return. As noted earlier, traditionally no findings from secondary research were fed back to participants, One project, the Personal Genome Project (PGP), gives its participants access to all the information it finds about them; however, it does not return findings from secondary researchers who use PGP data (PGP, 2011).

More recently studies are establishing a committee that will be in place at all times, ready to respond if a researcher discovers information that needs to be fed back. Decisions are made by this committee and then carried out by the study, or group of studies. This option has been chosen by the Coriell Personalized Medicine Collaboration (CPMC, 2011) and the eMERGE Consortium. Fullerton et al have presented details of the deliberations of the eMERGE Consortium Return of Results Oversight Committee (RROC) (Fullerton *et al.*, 2012). Once the RROC has made a decision as to which findings are potentially returnable, each eMERGE member study then considered the recommendations in light of their own cohort.

Such a committee can oversee this issue for a stand-alone research study, or for an international consortium (Wallace, 2011). If there is to be one oversight committee for a group of studies, such as in the case of eMERGE, it is key that the studies are sufficiently similar. This is so that all members can contribute effectively to the deliberations and any decisions can be considered for implementation by all the members. Limiting the variables may assist the decision-making process.

A further consideration is who will be on the committee and how decisions will be made. The Coriell Personalized Medicine Collaboration's Informed Consent Oversight Board is comprised of medical professionals, scientists, ethicists and community members (CPMC, 2011). This reflects best practice to include a wide range of stakeholders on such panels. As the decisions made by such a committee could have a significant impact on individuals (choosing one finding over another), it is vital that it has the appropriate expertise. Also, the process by which the decisions are made (i.e., by consensus, based on which particular evidence) needs to be clear, transparent and defendable.

2.6.2. Process by which IGR findings will be given to participants

Evidence is limited, but it appears that studies are feeding back findings either themselves or providing that information to the participant's GP, who will then give it to the participant. An example of the latter is the LifeLines cohort study in The Netherlands (Stolk *et al.*, 2008). If the LifeLines researchers discover a finding that is validated and clinically significant, they will give that information to the participant via their GP. This model makes use of an existing clinical relationship to manage the return process. Of course, this scenario will be most effective in a setting where individuals have a relatively stable, on-going relationship with a GP. It also presupposes that the participant has consented to their GP being informed of findings and that the study has up-to-date information on the GP of every study member. Even in the case of LifeLines, "[s]pecific genetic information is not returned so that the individual can say that they have not had genetic testing, as this has implications related to insurance" (Wallace and Kent, 2011).

If the CLS decides to manage the process itself, it will be necessary to determine how participants will be helped to cope, both clinically and emotionally, with any findings given to

them. Participants in a recent study said they preferred that HRFs be given to them in a face-to-face setting, and "…generally wanted to receive the result from someone with medical knowledge and expertise, who could ensure the finding was followed-up appropriately: usually a GP or a specialist healthcare professional" (Opinion Leader, 2012), p. 5. How this process will be managed is a vital piece of this complex puzzle.

2.7. Legal implications

A full analysis of the legal implications of returning IGR findings is beyond the scope of this briefing document. Indeed, there is little evidence on which to base such an analysis. There are clear responsibilities in the clinical setting (i.e. the duty of care owed a patient by his or her doctor), but the responsibilities of researchers are much less clear. US commentators state that, at this point, there is an ethical obligation for researchers to consider returning findings, but not a clear legal one (Clayton and McGuire, 2012). There is no case law in the UK that specifically concerns researchers or research participants (Boddington *et al.*, 2011), but there has been some research into consent materials and their legality. There are various pieces of UK legislation and international conventions that together place requirements on researchers to seek consent to participate in research (Boddington *et al.*, 2011), but the legality of research consent forms is questionable and subject to many factors. For example, if a consent form did state that findings would be returned, especially if specific mutations were listed, a participant might see this as a clear commitment on behalf of the researcher. If that participant was then not notified of such a finding and suffered harm as a result, they might believe that the researcher had acted negligently. Proving such a claim would depend on many issues, including the proximity of the researcher to the participant and whether that relationship was close enough to establish a duty of care.[1] However, until an action is brought and decided in the UK, the legal responsibilities owed to participants by researchers regarding returning IGR findings are not clear.

2.8 Responsibilities of secondary researchers

A decision will also need to be made as to the responsibilities of secondary researchers who may be discovering findings through their work. Details of responsibilities will need to be disseminated to, and agreed by, researchers seeking access to the cohort data. This could perhaps be as part of data-access agreements. At this time there is no general requirement for researchers to return findings, although in some countries researchers are being asked, if there is an expectation of discovering findings in their research, to include plans for how they will be returned to individuals in study protocols (Australian Government, 2007; (National Human Genomic Research Institute Intramural Research Bioethics Core, 2010); (Tri-Council Policy Statement, 2010). But these examples are for researchers setting up their own studies, so there is a question of whether these plans can be relied on by the CLS. The CLS could consider asking researchers for similar information on a data-access application. However, it is supposed that the CLS does not have the ability to force researchers to comply and can only resort to measures such as withdrawing access to data in cases of non-compliance, if such cases can be confirmed. Resources would clearly be needed to establish non-compliance.

3. Current recommendations for the process of feeding back results

If a decision is taken to feed back results, several steps will need to be taken. Using a framework can assist the CLS to delineate the steps needed to decide how to return IGR findings. One framework, recently proposed by Wolf et al (Wolf *et al.*, 2012), is the CARR framework (clarify, analyse, re-identify, re-contact) is used here as a suggested guideline. A 'return of results' policy reflecting the decisions made by the CLS and the process that will be followed will need to be developed, disseminated, and monitored.

3.1. Clarifying which IGR findings to return

Criteria are needed to determine which research results will be fed back. Once those are set, a list can be drawn up of the findings that meet those criteria. There is general consensus that studies *should* return findings, if consented to by the participant, that: (1) are analytically valid; (2) point to the possibility of suffering from an established and serious health condition; and (3) are clinically actionable. Wolf *et al* also suggest that studies may, additionally, return findings that "…reveal an established and substantial risk of likely health or reproductive importance or personal utility…" to a participant, as from their perspective they might benefit from knowing that information. Findings that are not likely to benefit the participant should not be returned (Wolf *et al.*, 2012) p. 18. Some examples of findings from the first two categories will be discussed in this paper. Reaching a consensus may be a difficult process, as opinions differ (Green *et al.*, 2012); care will need to be taken in choosing who will make the decisions and using what evidence. For example, as Richards mentions, the CLS might decide to only return those findings that are already used in clinical practice (Richards, 2011), although this is a moving target and therefore warrants close attention. Regardless, the process itself should be a transparent one, as far as is possible, including how decisions were reached and why.

Once a decision is made on the findings to be returned, these should be made into a list for use by the Cohort and for dissemination to participants and researchers seeking access to Cohort data. The criteria that were used should be included. The decisions taken should be reviewed regularly in response to the changing science. The process chosen to return findings should be reviewed regularly to determine if it is fit for purpose.

3.2. Analysing IGR findings from secondary researchers

The team managing the Cohort will need to have a process in place to review the finding brought by researchers to determine if it should be given to the participant. This will include whether the finding is on the approved list, and if not, should it still be considered for return. Other questions may include how and by whom the result was found, by what scientific methods, in what population, and if it meets the criteria set by the CLS. Members of a responsible committee, CLS members or someone independent of the CLS may be tasked with this review.

Scholars recommend (Bledsoe *et al.*, 2012) that all findings be validated for quality control and assurance purposes to ensure, for example, that approved methodologies have been used and that the finding is attributed to the correct person. This process naturally has resource implications; the CLS will need to decide if and how validation will take place.

Options here include the CLS requiring researchers to validate their own findings or detailing this job to an independent laboratory of a second opinion.

3.3. Re-identifying the cohort member

This entails re-identifying the individual participant so that they can be contacted. This will no doubt be done by CLS staff. The CLS will need a definitive record for each participant indicating whether or not they wish to be re-contacted with IGR information. This will help ensure that only those who want to know are contacted. However, if the finding is serious enough, the CLS will need to make a decision as to whether to override the participant's desire to not know. This will be based on whether the benefit to the individual outweighs, in CLS' opinion, the potential distress of the contact. Again, as in the validation step, it is vital that the correct person is identified for re-contact.

3.4. Re-contacting the cohort member

The re-contact process will need to be determined by the CLS and disseminated to cohort members. One method will be to contact them either directly or in writing, followed by discussion between the participant and the CLS, and then referral to a genetic counselling service. Another option is to give information on findings to the participant's general practitioner, as discussed earlier, who could then deal with the issue within a clinical setting. Wolf *et al* recommend that "[f]indings should be returned in a form that is understandable to the [participant] and useful to a physician or other clinician, such as a genetic counsellor…" (Wolf *et al.*, 2012) p. 19. Making the result 'understandable' is key, as explaining genetic risk to lay persons can present difficulties (Klitzman, 2010). It is important that additional distress is not caused due to clumsy communication.

3.5 Policy Development

A policy statement outlining what will or will not happen will need to be prepared, agreed, disseminated and monitored for effectiveness. The statement should include:

- A roster of the genetic variants and other information that will be considered for feeding back to participants;
- How they were chosen (by whom and using what criteria);
- The process that will be used in the future to return IGR findings; and
- The responsibilities of CLS and those of secondary researchers

It is important that the agreed process is monitored, as well as reviewed periodically, to see if it is being carried out effectively. Ongoing monitoring will help the CLS to know, for example, how often it is occurring, whether there are some findings that are appearing more frequently than others, and whether the list of findings should be updated or changed. In addition, there should be some form of gauging how the process is working for cohort members. This could, for example, include instituting some means, whether as research or consultation, to explore the experience with those who had had an IGR finding given to them. The information gained could feed into the ongoing review of the policy.

4. Research that might produce results

There is much debate regarding which data should be returned, but some evidence is emerging and groups are making suggestions based on their own experience (Fullerton *et al*., 2012). Berg et al suggest using a 'binning' system to categorise genetic variants that might be considered for return, ranging from those that should be returned, those that are debatable and those that should not (Berg *et al*., 2011). Examples fall into these bins or categories based on how well they meet agreed criteria for clinical validity, or "…how well the [genetic information] predicts the presence or absence of the phenotype, clinical disease or predisposition" (HGSG, 2012 p.42) and clinical utility, whether having the genetic information "…will inform patient management and result in an improved clinical outcome" (HGSG, 2012, p. 43). As discussed earlier, commentators recommend that results that meet both criteria should be considered for return to participants. However, if the requirement of clinical utility is not met, this is considered borderline and returning the information should be considered on a case-by-case basis.

Any result given to a participant that indicates they have, or are predisposed to having, a condition will need to be confirmed in a clinical setting. The examples in category 1 have shown sufficient evidence to suggest that, even with any extra physical or mental distress that is caused by additional testing, or any additional costs, an individual should benefit clinically from being given the information. It is unclear, for conditions in category 2, whether the additional cost and distress caused by additional testing is warranted. The cohort will need to remember that some participants will want to have any 'borderline' findings given back to them so that they can make their own decisions on how to act on that information.

Berg et al suggest that there are currently only about 100 category 1 variants, while there are thousands in category 3; others suggest millions might be a better estimate for the latter category. The data falling into these categories will necessarily change as knowledge changes. The results that will be fed back will be different for each study and each should be examined on a case-by-case basis.

Table 4 (adapted from Berg *et al*., 2011) shows the three categories, their criteria, the suggested action and examples from the literature. Those marked have either direct or suggested applicability to 1958BC participants and will be discussed in more detail.

Table 4. Categories of findings and examples

Cat.	Definition	Suggested Action	Examples from the Literature
1	Direct clinical utility based on current medical literature	Should be fed back routinely	Neurofibromatosis Type 1, Lynch Syndrome, *BRCA1/2* (Berg *et al.*, 2011) Turner Syndrome, Klinefelter Syndrome (Fullerton *et al.*, 2012) Familial hypercholesterolaemia*
2	Clinical validity but no strongly actionable implications	May be fed back, as the information may be useful to or	PGx variants and common risk SNPs, *APOE4* allele, Huntington (Berg *et al.*, 2011)

		wanted by individuals	*HFE* haemochromatosis*, Factor V Leiden, (Fullerton *et al.*, 2012) *m.1555A>G* (predisposition to aminoglycoside-induced deafness)*, Coeliac disease*
3	Has not been strongly linked to a phenotype, clinical outcome or intervention	Should not be fed back	

*Of relevance to the 1958BC and discussed further below.

4.1. Recommended returnable findings

The variants in genes or loci in Bin 1 (Table 4) have shown clinical utility and are actionable. As noted, the current number of these is relatively small, although the list will expand with further research. One particular example, familial hypercholesterolaemia, may be significant for members of the 1958BC cohort.

4.1.1. Familial hypercholesterolaemia

Familial hypercholesterolaemia (FH) is a monogenic disorder that carries a high risk of premature coronary heart disease (>50% risk in men by age 50 and >30% in women by age 60); the estimated prevalence in the UK population is 1 in 500 (Wierzbicki *et al.*, 2008). The disease is treatable by statins and lifestyle changes, but is currently under diagnosed. UK NICE guidelines recommend that a suspected case of FH should be confirmed through genetic testing to confirm the diagnosis and to trigger cascade testing of relatives (Wierzbicki *et al.*, 2008). One expert has estimated that in a cohort of approximately 10,000, approximately 50 people will be show a potentially actionable mutation predisposing heart attack through exome sequencing; a majority of these will be FH. Approximately 30-40 of these individuals will have their diagnosis confirmed, triggering cascade testing leading to 100-150 others to be approached regarding treatment.[2] For a cohort the size of the 1958BC, this could mean that approximately 37 people will have the mutation (again with the majority being FH), 20-30 of whom would have their diagnosis confirmed, with 75-100 further family members to be tested. Cascade testing is seen as a cost-effective means of finding undiagnosed FH patients (Humphries, 2011).

There are major benefits to this scenario. If an FH mutation is found by a secondary researcher and it is decided to disclose this to the cohort participant, there is a recognised clinical procedure and effective treatments. Making use of this established process should not require additional resources from the cohort. Cascade screening for relatives, while recommended by NICE, is currently funded in Scotland, Wales and Northern Ireland, but not in England (Humphries, 2011). This may mean that if a participant in Bristol is diagnosed, there could be no money for testing in their relatives' local authority. However, this should not necessarily preclude feeding back an FH mutation status to the participant.

4.2. Findings of Questionable Returnability

Bin 2 holds those variants that have clinical validity but do not have clear clinical interventions. However, Berg et al (Berg *et al.*, 2011) note that individuals might still want to know that information, raising the question of personal utility. For example, knowing one's carrier status for a condition might be useful for those of reproductive age, but not for others. Therefore, decisions around feeding back these results will be more complex. Three examples that are applicable to the 1958BC participants illustrate this point. Examples 4.2.2 and 4.2.3 are from studies using 1958BC data.

4.2.1. *HFE*-associated hereditary haemochromatosis

HFE-associated hereditary haemochromatosis (HH) is the most common type of inherited iron overload disorder. The p.C282Y mutation in the *HFE* gene is the most common form and highly prevalent in those of Northern European descent. HH most often occurs when a

person has two copies of the C282Y mutation (homozygous). Those with HH suffer from, for example, cirrhosis, liver malfunction, and hepatocellular carcinoma. Iron depletion by phlebotomy is the main treatment and is highly effective (Alexander and Kowdley, 2009).

The eMERGE RROC considered whether or not to feed back HH information to its participants (Fullerton *et al.*, 2012). They noted in their discussions that the penetrance of the disease is low in men and even lower in women, and that a minority of people develop clinical disease. But as there is an effective treatment, the RROC weighed in favour of contacting homozygous men. However, when each of the eMERGE member studies looked at the men in their own studies who could be contacted, for various reasons, none of them were. For example, the Marshfield Clinic found six homozygous men, three of whom had not previously been diagnosed with HH. When they examined these cases more closely, it appeared those who had not already been diagnosed showed no symptoms that suggested they had HH. Based on discussions, careful review of existing consents that stated that no results would be returned, and reviews of existing guidelines, the decision was made not to contact those men (Fullerton *et al.*, 2012).

This decision is completely understandable, yet it does raise the question of whether those individuals, if asked, might have wanted that information regardless of the lack of clinical symptoms. Some might, as they would be aware of the possibilities and could act accordingly if symptoms appeared. Others might not wish to know, as it could cause unnecessary distress if symptoms never appeared. This case highlights the difficulty with 'borderline' results and a clear and transparent process by which decisions are reached.

4.2.2. Predisposing Aminoglycoside-Induced Deafness

The m.1555A>G mutation has been shown to predispose individuals to be hypersensitive to aminoglycosides, which are used to treat infections such as multi-resistant strains of *Escherichia coli* (Babiker, 2007). Treatment with drug levels in the therapeutic range can cause profound and permanent deafness. As well, it is believed that the mutation may cause late-onset hearing loss even in those not exposed to aminoglycosides. Rahman et al genotyped 1958BC samples and compared mutation data with hearing outcome data collected from participants (Rahman *et al.*, 2012). They found that 1 in 385 of the 1958BC members had this mutation, but that their hearing was no different than that of the general population. However, the authors did recommend that an individual with this mutation have genetic testing for m.1555A>G prior to aminoglycoside administration in order to prevent inevitable hearing loss.

There is an obvious benefit for individuals in knowing that they have this mutation, if it can prevent unnecessary hearing loss. But there are opposing views and questions raised. For example, how often will an individual will be treated with an aminoglycoside? As it is a relatively inexpensive antimicrobial agent, Rahman et al suggest that their use will increase, but this is not confirmed. Others have questioned the practicality of the genetic test. Rourke notes that it often takes time for genetic tests to be run and as, "[a]minoglycoside treatment is started most commonly in the acute setting,…withholding treatment for two or three days for the results of a genetic screening test would not be in the best interests of the patient" (Rourke, 2007), p. 952. Others note that as the penetrance of this mutation is approximately of the same magnitude as the risks from other commonly used antibiotics, setting aside an effective treatment because a very few would suffer may not be appropriate (Babiker, 2007).

This example presents many different questions and viewpoints. The benefit to the individual might be clear, yet the impact on the NHS, and therefore its users, is less clear.

4.2.3. Coeliac Disease

Coeliac Disease (CD) is an autoimmune disease; eating gluten triggers a reaction that damages the lining of the small intestine.(Coeliac UK) HLA genes have been identified, but 60% of cases are attributable to an unknown number of non-HLA genes (Gutierrez-Achury *et al.*, 2011).

Trynka et al used 1958BC data, with other datasets, to identify common and rare variant association signals in CD (Trynka *et al.*, 2011). Their work, in their view, is the first to look comprehensively for all known risk loci for a trait, in order to answer questions regarding heritability and determine which genes are causal. But they acknowledge that more work is needed. “Although we localized signals at many loci,…only more detailed functional studies…will show precisely which gene variants might be causal” (Trynka *et al.*, 2011), p. 1200. In addition, experts suggest that a proper characterisation of the phenotype is still needed, as CD has a broad number of symptoms that overlap with other autoimmune diseases (Gutierrez-Achury *et al.*, 2011).

In this case, the CLS may decide not to give out this information. Yet, an individual can readily obtain a personal diagnosis for coeliac. MyCeliacID is a ‘do-it-yourself’ test that purports to tell you your risk of having CD (Prometheus Labs, 2012); 23andMe gives out data on four report markers for coeliac disease (23andMe, 2012). This is an interesting example because it draws a comparison between information a research study may be willing to disclose and what is available elsewhere.[3] A policy not to feed back such results may prompt participants to question how this decision was reached.

5. Summary

In summary, the decision process regarding the return of IGRs is a complex one. It has been the aim of this paper to outline some of the issues and provide some examples of how other cohort studies are approaching this issue. As well, it has shown that research studies currently using 1958BC data are identifying genetic variants that might warrant considering feeding them back to participants. However, the CLS will first need to decide whether or not it will return IGR findings at all. This decision making process should consider the following questions:

- Does the original consent language already state that findings will be returned to participants?
- If not, is returning findings a new use of participants' data that warrants re-consent?
- Is re-consent feasible, based on, for example, the size of the cohort and the ability to re-contact the vast majority of those who originally gave consent?
- What method of re-consent might be used (i.e., seeking new consent, opt-out or notification only)?
- Does the expected cost of re-consent rule out re-consent?
- If not, where can the funding for re-consent be found?
- How will the CLS manage the process by which findings are given to participants?
- Will the CLS wish to create its own oversight committee, or join others to form a joint committee?
- Who will be on any oversight committee and how will it be managed, governed and monitored?
- How will it monitor the process of returning findings to participants?
- What are the legal implications for the CLS related to returning findings?
- What are the responsibilities of secondary researchers who might discover findings?
- Who will be responsible for any costs associated with returning results?

If it is decided to return IGR findings, the next step will be to define the process by which this will happen. This will involve examining the following questions:

- What criteria will the CLS use to decide which findings?
- How will the decisions be made, by whom and through what process?
- Based on these criteria, which findings will be returned to 1958BC participants?
- What will be included in any policy statement?
- How will this policy be disseminated to stakeholders such as participants, researchers who may wish to access 1958BC data, the public, funders and the CLS staff?
- How will findings presented by a secondary researcher be validated?
- How and by whom will participants be contacted?
- How will the CLS ensure that they are going to contact the correct person and that person wishes to be contacted with such information?
- How and by whom will the CLS actually deliver the information to the participant?
- How will the CLS ensure that the participant is being given adequate emotional and medical support through the return process?
- How will the CLS ensure that their process for returning IGR findings is working?

In conclusion, the return of IGRs is a complicated and value-laden issue. The CLS needs to act, as far as is possible, according to the best available evidence, using the best available

minds, in an open, transparent and defensible manner, in order to ensure the best interests of its cohort members are preserved.

References [suggested references for further reading are in bold]

23andMe. (2012), Get tested to learn what your genetics say about: Celiac Disease, [accessed 02May 2012], available at https://www.23andme.com/health/Celiac-Disease/.

Alexander, J. and Kowdley, K.V. (2009), 'HFE-associated hereditary hemochromatosis', *Genet Med*, 11: 5, 307-313.

Australian Government (2007), National Statement on Ethical Conduct in Human Research.

Babiker, Z.O.E. (2007), Aminoglycosides: putting risk in perspective, [accessed 15 May 2012], available at: http://www.bmj.com/rapid-response/2011/11/01/aminoglycosides-putting-risk-perspective.

Berg, J.S., Khoury, M.J. and Evans, J.P. (2011), 'Deploying whole genome sequencing in clinical practice and public health: Meeting the challenge one bin at a time', *Genetics in Medicine*, 13: 6, 499-504 410.1097/GIM.1090b1013e318220aaba.

Bledsoe, M.J., Grizzle, W.E., Clark, B.J. and Zeps, N. (2012), 'Practical implementation issues and challenges for biobanks in the return of individual research results', *Genet Med*, 14: 4, 478-483.

Boddington, P., Curren, L., Kaye, J., Kanellopoulou, N., Melham, K., Gowans, H. and Hawkins, N. (2011), 'Consent Forms in Genomics: The Difference between Law and Practice', *European Journal of Health Law*, 18: 5, 491-519.

Bollinger, J.M., Scott, J., Dvoskin, R. and Kaufman, D. (2012), 'Public preferences regarding the return of individual genetic research results: findings from a qualitative focus group study', *Genet Med*, 14: 4, 451-457.

Christensen, K.D., Roberts, J.S., Shalowitz, D.I., Everett, J.N., Kim, S.Y.H., Raskin, L. and Gruber, S.B. (2011), 'Disclosing Individual CDKN2A Research Results to Melanoma Survivors: Interest, Impact, and Demands on Researchers', *Cancer Epidemiology Biomarkers & Prevention*, 20: 3, 522-529.

Clayton, E.W. and McGuire, A.L. (2012), 'The legal risks of returning results of genomics research', *Genet Med*, 14: 4, 473-477.

Coeliac UK. 2012. What is coeliac disease? [accessed 02 May 2012], available at http://www.coeliac.org.uk/coeliac-disease/what-is-coeliac-disease.

CPMC (2011), Coriell Personalized Medicine Collaborative, [accessed 20 April 2011], available at: http://cpmc.coriell.org/.

Fullerton, S.M., Wolf, W.A., Brothers, K.B., Clayton, E.W., Crawford, D.C., Denny, J.C., Greenland, P., Koenig, B.A., Leppig, K.A., Lindor, N.M., McCarty, C.A., McGuire, A.L., McPeek Hinz, E.R., Mirel, D.B., Ramos, E.M., Ritchie, M.D., Smith, M.E., Waudby, C.J., Burke, W. and Jarvik, G.P. (2012), 'Return of individual research results from genome-wide association studies: experience of the Electronic Medical Records and Genomics (eMERGE) Network', *Genet Med*, 14: 4, 424-431.

Green, R.C., Berg, J.S., Berry, G.T., Biesecker, L.G., Dimmock, D.P., Evans, J.P., Grody, W.W., Hegde, M.R., Kalia, S., Korf, B.R., Krantz, I., McGuire, A.L., Miller, D.T., Murray, M.F., Nussbaum, R.L., Plon, S.E., Rehm, H.L. and Jacob, H.J. (2012), 'Exploring concordance

and discordance for return of incidental findings from clinical sequencing', *Genet Med*, 14: 4, 405-410.

Gutierrez-Achury, J., Coutinho de Almeida, R. and Wijmenga, C. (2011), 'Shared genetics in coeliac disease and other immune-mediated diseases', *Journal of Internal Medicine*, 269: 6, 591-603.

Humphries, S.E. (2011), 'Guidelines for the identification and management of patients with familial hypercholesterolaemia (FH): Are we coming to a consensus?', *Atherosclerosis Supplements*, 12: 2, 217-220.

Klitzman, R. (2010), 'Misunderstandings Concerning Genetics Among Patients Confronting Genetic Disease', *Journal of Genetic Counseling*, 19: 5, 430-446.

Knoppers, B.M. and Kharaboyan, L. (2009), '"Deconstructing" biobank communication of results', *SCRIPTed*, 6: 3, 677.

Ludman, E.J., Fullerton, S.M., Spangler, L., Trinidad, S.B., Fujii, M.M., Jarvik, G.P., Larson, E.B. and Burke, W. (2010), 'Glad You Asked: Participants' Opinions Of Re-Consent for dbGap Data Submission', *Journal of Empirical Research on Human Research Ethics: An International Journal*, 5: 3, 9-16.

National Human Genomic Research Institute Intramural Research Bioethics Core (2010), Points to Consider in the Transition Toward Whole-Genome Sequencing in Human Subjects Research, [accessed 20 April 2011], available at: http://www.genome.gov/Pages/Research/Intramural/IRB/WES-WGS_Points_to_Consider.pdf.

Opinion Leader (2012), Assessing Public Attitudes to Health Related Findings in Research, [accessed 18 May 2012], available at: https://www.wellcometrustevents.org/WELLCOME/media/uploaded/EVWELLCOME/event_124/WT%20MRC%20HRF%20report%20%28website%20version%29.pdf.

PGP (2011), Personal Genome Project Consent Form, [accessed 20 April 2011], available at: http://www.personalgenomes.org/consent/PGP_Consent_Approved_02222011.pdf.

Prometheus Laboratories Inc (2012), MyCeliacID, [accessed 2 May 2012], available at: https://myceliacid.com/.

Rahman, S., Ecob, R., Costello, H., Sweeney, M.G., Duncan, A.J., Pearce, K., Strachan, D., Forge, A., Davis, A. and Bitner-Glindzicz, M. (2012), 'Hearing in 44–45 year olds with m.1555A>G, a genetic mutation predisposing to aminoglycoside-induced deafness: a population based cohort study', *BMJ Open*, 2: 1.

Richards, M. (2011), 'An ethical review of the use of functional MRI, CLS Working Paper 2011/4', London: Centre for Longitudinal Studies.

Rourke, T. (2007), 'How practical is this genetic screening test?', *BMJ*, 335: 7627, 952-952.
Steinsbekk, K.S. and Solberg, B. (2011), 'Biobanks—When is Re-consent Necessary?', *Public Health Ethics*.

Stolk, R., Rosmalen, J., Postma, D., de Boer, R., Navis, G., Slaets, J., Ormel, J. and Wolffenbuttel, B. (2008), 'Universal risk factors for multifactorial diseases', *European Journal of Epidemiology*, 23: 1, 67-74.

Tri-Council Policy Statement (2010), Ethical Conduct for Research Involving Humans, [accessed 14 March 2010], available at: http://www.pre.ethics.gc.ca/pdf/eng/tcps2/TCPS_2_FINAL_Web.pdf.

Trinidad, S.B., Fullerton, S.M., Ludman, E.J., Jarvik, G.P., Larson, E.B. and Burke, W. (2011), 'Research Practice and Participant Preferences: The Growing Gulf', *Science*, 331: 6015, 287-288.

Trynka, G., Hunt, K.A., Bockett, N.A., Romanos, J., Mistry, V., Szperl, A., Bakker, S.F., Bardella, M.T., Bhaw-Rosun, L., Castillejo, G., de la Concha, E.G., de Almeida, R.C., Dias, K.-R.M., van Diemen, C.C., Dubois, P.C.A., Duerr, R.H., Edkins, S., Franke, L., Fransen, K., Gutierrez, J., Heap, G.A.R., Hrdlickova, B., Hunt, S., Izurieta, L.P., Izzo, V., Joosten, L.A.B., Langford, C., Mazzilli, M.C., Mein, C.A., Midah, V., Mitrovic, M., Mora, B., Morelli, M., Nutland, S., Nunez, C., Onengut-Gumuscu, S., Pearce, K., Platteel, M., Polanco, I., Potter, S., Ribes-Koninckx, C., Ricano-Ponce, I., Rich, S.S., Rybak, A., Santiago, J.L., Senapati, S., Sood, A., Szajewska, H., Troncone, R., Varade, J., Wallace, C., Wolters, V.M., Zhernakova, A., Thelma, B.K., Cukrowska, B., Urcelay, E., Bilbao, J.R., Mearin, M.L., Barisani, D., Barrett, J.C., Plagnol, V., Deloukas, P., Wijmenga, C. and van Heel, D.A. (2011), 'Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease', *Nat Genet*, 43: 12, 1193-1201.

Wallace, S.E. (2011), 'The Needle in the Haystack: International Consortia and the Return of Individual Research Results', *The Journal of Law, Medicine & Ethics*, 39: 4, 631-639.

Wallace, S.E. and Kent, A. (2011), 'Population biobanks and returning individual research results - mission impossible or new directions?', *Human Genetics*, 130: 3, 393-401.

Wierzbicki, A.S., Humphries, S.E. and Minhas, R. (2008), 'Familial hypercholesterolaemia: summary of NICE guidance', *BMJ*, 337.

Wolf, S.M., Crock, B.N., Van Ness, B., Lawrenz, F., Kahn, J.P., Beskow, L.M., Cho, M.K., Christman, M.F., Green, R.C., Hall, R., Illes, J., Keane, M., Knoppers, B.M., Koenig, B.A., Kohane, I.S., LeRoy, B., Maschke, K.J., McGeveran, W., Ossorio, P., Parker, L.S., Petersen, G.M., Richardson, H.S., Scott, J.A., Terry, S.F., Wilfond, B.S. and Wolf, W.A. (2012), 'Managing incidental findings and research results in genomic research involving biobanks and archived data sets', *Genet Med*, 14: 4, 361-384.

Notes:

1. Dr Susan E Wallace is Lecturer in Population and Public Health Sciences, Department of Health Sciences, University of Leicester, sew40@le.ac.uk

2. This is based on research being conducted by colleagues at the University of Edinburgh and is as yet unpublished.

3. Personal communication.

4. My thanks to Neil Walker for raising this question.