



LJMU Research Online

Brett, CE and Deary, IJ

Realising health data linkage from a researcher's perspective: following up the 6-Day Sample of the Scottish Mental Survey 1947

<http://researchonline.ljmu.ac.uk/id/eprint/2221/>

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Brett, CE and Deary, IJ (2014) Realising health data linkage from a researcher's perspective: following up the 6-Day Sample of the Scottish Mental Survey 1947. *Longitudinal and Life Course Studies*, 5 (3). pp. 283-298. ISSN 1757-9597

LJMU has developed [LJMU Research Online](#) for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

<http://researchonline.ljmu.ac.uk/>

Realising health data linkage from a researcher's perspective: following up the 6-Day Sample of the Scottish Mental Survey 1947

Caroline E Brett Department of Psychology, University of Edinburgh
caroline.brett@ed.ac.uk

Ian J Deary Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology,
University of Edinburgh

(Received August 2013 Revised July 2014)

<http://dx.doi.org/10.14301/llcs.v5i3.266>

Abstract

Health and wellbeing in old age are influenced by genetic, environmental and social factors throughout the life course. At present, few longitudinal studies offer information from childhood through to old age. Data linkage between multiple sources of health data enhances the value of existing longitudinal data. Regulations governing access to personal data for health research exist to protect the privacy and confidentiality of data on behalf of the individual. This paper outlines the process of obtaining permission for data linkage from a researchers' perspective, using a case study which offers an unusual opportunity to understand life course influences such as socio-economic status, childhood deprivation and measured intelligence on health and wellbeing in old age in an entire year-of-birth population. The Scottish Mental Survey 1947 (SMS1947, n = 70,805) has childhood intelligence data from individuals born in 1936 and attending schools in Scotland in June 1947. Representative sub-groups of the SMS1947 provided additional sociological information. The 6-Day Sample (n = 1,208), born on 6 days of 1936, were followed up for 16 years to age 27. Their younger siblings also took an intelligence test and were followed up for several years. Our team's planned research on the SMS1947 falls into two distinct parts. The first is a revival of the 6-Day Sample study involving tracing Sample members and inviting survivors to a follow-up study. The second part aims to carry out linkage between existing data on the SMS1947, its sub-groups, and the younger siblings, and morbidity and mortality data from central databases in Scotland and in England and Wales. We conclude by offering some recommendations for simplifying the process of obtaining permission to access linked health data, and place these into the context of the shifting landscape of data linkage in the UK and beyond.

Keywords: Scottish Mental Survey; intelligence; ageing; health; wellbeing; life course epidemiology; data linkage; health outcomes

Introduction

There is a growing acknowledgement amongst researchers in epidemiology and public health of the influence of biological, physical and social exposures across the life course on subsequent health and wellbeing (Blane, Netuveli & Stone, 2007; Foresight Mental Capacity and Wellbeing Project, 2008; Kuh, Ben-Shlomo, Lynch, Hallqvist &

Power, 2003; Kuh, Cooper, Hardy, Richards & Ben-Shlomo, 2014; Lynch & Davey Smith, 2005; Richter & Blane, 2013). Life course research provides an inter-disciplinary framework for understanding how genetic as well as environmental factors such as socio-economic status, occupational hazards and childhood deprivation influence medical and social

inequalities in health and mortality. Life course epidemiology relies heavily on the use of longitudinal studies, which enable researchers to track the health and social trajectories of individuals over time.

Cohort studies are particularly valuable if they follow the same individuals from early life into adulthood (Pearson, 2011). They can reduce confounding by factors such as age or geographical location. The British birth cohorts, for example, offer a wealth of information from birth or early years through adulthood and into early old age, and have helped to address key questions on health inequalities, mental wellbeing, health behaviours, and genetic influences on health and ageing (Cooper et al., 2012; Kuh et al., 2012; Pearson, 2011; Power & Elliott, 2005; Wadsworth, Kuh, Richards & Hardy, 2006).

The oldest British birth cohort, born in 1946, reached retirement age in 2011 (Wadsworth et al., 2006). This restricts, for the moment, the usefulness of the British cohorts in answering questions about health in older age. A number of longitudinal cohort studies exist specifically to investigate ageing processes. A recent systematic review identified 67 longitudinal studies of older adults across all continents (Seematter-Bagnoud & Santos-Eggimann, 2006). Most participants were aged 50 or above at the time of recruitment, and although the length of follow-up varied from 2 to more than 30 years, the majority followed participants up for less than 10 years. Both factors limit the usefulness of these studies in investigating life course influences on ageing. In the UK, longitudinal studies of ageing include the Whitehall II study of retired civil servants (Marmot & Brunner, 2004), and the English Longitudinal Study of Ageing (Stephens, Breeze, Banks & Nazroo, 2012). Both have produced many findings on health and wellbeing from the latter stages of working life into retirement and old age. However, they include minimal information from childhood and early adulthood, which are key periods of transition and exposure to social and environmental influences (Foresight Mental Capacity and Wellbeing Project, 2008).

Another valuable resource to life course epidemiology is population-based data collected by government or military organisations. These sometimes contain detailed medical, psychological and cognitive records on large numbers of

individuals. For example, many epidemiologists have made use of Swedish conscription data, which is not only detailed but involves over 1 million young men followed for decades (Batty et al., 2007, 2009). The Vietnam Experience Study, although smaller in number, has been used to investigate determinants of health and mortality (Batty et al., 2008; Weiss, Gale, Batty & Deary, 2009). Udjus (1964) conducted innovative research on the relationships between physical growth and intelligence, education and family size using data from Norwegian conscripts (Tanner, 1966; Udjus 1964). However, military-based studies exclude women, and individuals who do not meet strict medical criteria. By contrast, data on cognitive function and educational attainment are available for much of the population through schools and examination boards, and can greatly contribute to longitudinal studies (e.g. Booth et al., in press; Calvin, Fernandes, Smith, Visscher & Deary, 2010; Deary, Strand, Smith & Fernandes, 2007).

Data linkage to enhance longitudinal research

The majority of longitudinal studies still rely on data provided by the participants themselves, either by self-report or the collection of objective health measures such as biomarkers. However, the value of these data is reduced by attrition and the inaccuracies of self-report data. Maintaining regular contact with longitudinal study participants is costly in terms of time, effort and resources. Researchers are, therefore, increasingly turning to alternative data sources. One such source is the computerised administrative health (and non-health) data which most developed countries now collect routinely. The primary purpose of these data is to enhance planning and policy decision-making by providing detailed information on the population's health and service use. This type of health data commonly comprises death certification, hospital admissions, hospital consultations and prescription records, with increasing coverage of consultation at general practitioners, disability assessment and sickness absence. Their value easily extends to medical research, providing extensive population-based information on the aetiology of disease, its course and outcome. In longitudinal cohort studies, linkage between research data and routinely-collected health records can be used to verify the quality and accuracy of self-report data such as medical history. However, administrative health data collected at point of service use is limited in scope, including

only individuals who consult health services. Furthermore, unlike some military or conscript data, health records generally contain scant information on health behaviours or function.

The research value of administrative health data can be further enhanced when data from different sources relating to the same individuals are combined. Data linkage of this kind is not a new concept (Dunn, 1946). The linkage of health data with non-health data such as social care, benefits, crime or census records, has a multitude of benefits, not only to medical research but also in terms of public health, cost efficiency and administrative accuracy (Holman et al., 2008, Scottish Government, 2010). The use of large, multiple, population-based administrative data sets conserves patient privacy by using a single identifier, thereby reducing the need for personal identifiers such as name and date of birth (Holman et al., 2008). Due to the organisational and collaborative effort required, very few such data sets exist internationally, and most by necessity tend to exist at a regional rather than country-wide level. Linked data sets currently exist in Manitoba, Canada; British Columbia, Canada; Rochester, Minnesota; Oxford, England; Scotland; Western Australia; and New South Wales, Australia (Roos, Menec & Currie, 2004; Scottish Government, 2010). One exception is Finland, which holds personal data on its population in large, linked data sets (Gissler & Haukka, 2004). In the UK, the Longitudinal Study – England and Wales (Goldring & Newman, 2010), the Northern Ireland Longitudinal Study (O'Reilly, Rosato, Catney, Johnston & Brolly, 2012) and the Scottish Longitudinal Study (Boyle et al., 2009) have successfully linked health and non-health data, including Census data, on a representative sample of these countries' populations. These linked data sets represent excellent examples of collaborative working and have enabled researchers and policy makers to address important research questions around health inequalities that would otherwise have been impossible to answer.

Rapid technological advances in recent years have made it possible to collect, store and analyse vast amounts of personal data and this has led to a worldwide debate around issues of privacy and confidentiality (Regidor, 2004). Data protection legislation was introduced in countries throughout the world, with the explicit aim of protecting the individual's right to privacy while recognising the

need to collect and store personal data (e.g. European Union (EU) Data Protection Directive 95/46/EC, 1995; UK Data Protection Act, 1998). Health data collected at the point of use is subject to the common law duty of confidentiality. Patients use health services with the expectation that the information they are providing will be kept confidential, and registered clinicians are bound by the statutes and regulations of their professions. As a consequence, additional legislation governs the use of personal data relating specifically to health. In the UK, the government-commissioned Caldicott Report of 1997 highlighted six key principles and made 16 recommendations about the flow and handling of patient information within the NHS (The Caldicott Committee, 1997). Caldicott Guardians were appointed within each NHS organisation to uphold these principles. More recently, information governance frameworks incorporating legal rules, guidance and best practice were put in place to guard against the inappropriate use of patient information. In Finland, the Personal Data Act stipulates that health and social information can only be gathered by informed consent from the individual, with the exception of data collected for statistics and historical or scientific research (Gissler & Haukka, 2004).

The issue of consent

The majority of debate concerning the use of health data in research, centres around the issue of consent. Mason and Laurie (2010) argue that "it is undeniable that consent remains the primary policy device in legitimating medical research". Indeed, all medical and epidemiological research that requires the presence of individuals should be carried out in accordance with the ethical principles of research on human subjects laid out in the Declaration of Helsinki (1964) (Laurie & Sethi, 2011; Regidor, 2004). Research subjects must give voluntary consent, after having been informed of the purpose, methods and possible risks and benefits of the research study in question. Of course, in many epidemiological or longitudinal studies using health data, obtaining informed consent from each and every individual is not possible. This is particularly the case in studies with large numbers of subjects, where obtaining consent would be too costly and time-consuming, or where subjects have died, or in studies using historical data (Laurie & Sethi, 2011; Regidor, 2004). The requirement to obtain informed consent can also introduce bias, as those who do

not provide consent often systematically differ from those who do (Al-Shahi Salman et al., 2014; Kho, Duffett, Willison, Cook & Brouwers, 2009). Further issues arise in longitudinal research beginning in childhood, where consent needs to be updated once respondents reach an appropriate age. The EU Directive (1995) currently makes provision for circumstances in which obtaining informed consent may not be possible. Article 81 states that although consent must ideally be obtained for all research using personal data, exceptions could be made for research serving “exceptionally high public interests” and that “cannot possibly be carried out otherwise” (Directive 95/46/EC, 1995, Mansell, 2013).

The ethical and moral implications of the use of health data without consent have been discussed in detail by Regidor (2004). Regidor highlights the impact that differing interpretations of the EU Directive have on access to personal medical data in countries across the EU. Often, restrictions on access to data are placed by the data custodians themselves, with considerable variation within and between countries. Even where data sets are linked, researchers may be required to apply for separate approval from each data custodian, as occurs in the Western Australia Data Linkage System (Holman et al., 2008), or for further approvals where the data set is particularly large or contains particularly sensitive health information, as with the Finnish health and social welfare registers (Gissler & Haukka, 2004). In the United States, legislation varies by state, rendering it almost impossible to conduct population level health research and introducing the possibility that key demographic or social groups are under-represented in medical research (Melton, 1997; Regidor, 2004). Progress, some argue, is informed less by the benefits of such research and more by the perceived risks (Flowers & Ferguson, 2010; Laurie & Sethi, 2011; Regidor, 2004). Certainly, the administrative burden of complying with the requirements of research regulation and management, causes delay and wastage in medical research (Al-Shahi Salman et al., 2014).

Data linkage in the UK

The situation in the UK appears, at first, to be more promising. In recent years there has been an increasing acknowledgement within the government of the value of health data in life course research. The UK's current Prime Minister

argued for the opening up of patient data for research (Cabinet Office and Prime Minister's Office, 2011; Walker, 2011). There is a recognised need for systems to enable smooth, safe and secure linkage between research studies and health and other records. Recent initiatives such as the Medical Research Council's funding of eHealth Centres, the Economic and Social Research Council's funding of four Administrative Data Research Centres across the UK, and infrastructure initiatives such as the Scottish Informatics Programme (SHIP), the NIHR Research Capability Programme in England, and the Welsh Secure Anonymised Information Linkage System (SAIL), are all working towards the availability of electronic health records for research purposes within a secure and ethically-approved environment.

Clearly then, there is potential value in data linkage, there are databases and information structures in place to support it, and there is encouragement to do it from the scientific community and the government. However, the reality of gaining permission to carry out data linkage is often complex, with inconsistencies in the interpretation of legislation and the restrictions placed on access to different types of health data (Laurie & Sethi, 2011; Regidor, 2004). In this paper, we present a case study of the processes involved in our obtaining permission to link health data, held within the UK, with existing life course data from the Scottish Mental Survey 1947. We will then make recommendations based on our experience and discuss the shifting landscape of data linkage in the UK and what it might mean for researchers.

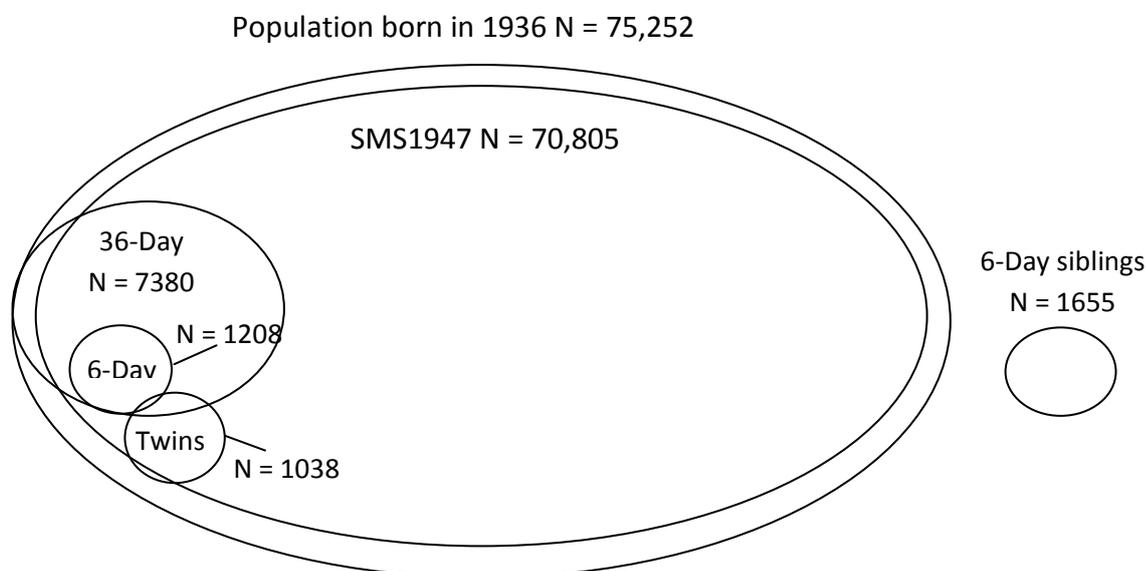
The Scottish Mental Survey 1947 and its proposed research uses

The Scottish Mental Survey 1947 (SMS1947) took place on 4th June 1947 (Deary, Whalley & Starr, 2009; Mental Survey Committee, 1949). It tested the intelligence of almost all children born in 1936 and attending Scottish schools. It applied a version of the Moray House Test No. 12 to 70,805 children. It was organised by the Scottish Council for Research in Education. Three subsamples were tested afterwards to collect more information. The 7,380 children born on the first three days of each month had a four-page 'Sociological Schedule' filled in; they were named the 36-Day Sample (Mental Survey Committee, 1953, 1958). The Sociological Schedule data were also collected on all 542 twin

pairs born in 1936. The 1,208 children born on the first day of the even-numbered months had, in addition to the Sociological Schedule, an individual Stanford-Binet IQ test and home visits and interviews almost every year until they were 27. They were named the 6-Day Sample (MacPherson,

1958; Maxwell, 1961, 1969). Their younger siblings had the Stanford-Binet IQ test applied as they reached age 11. Eventually, 1,554 of these siblings were tested. Figure 1 illustrates the relationships between the sub-groups of the SMS1947.

Figure 1. Venn diagram of the Scottish Mental Survey 1947 and its sub-groups



The 6-Day Sample offers a rich and detailed source of information on childhood and early adulthood in a representative sample of the Scottish population born in 1936. The existing data has already been used to answer research questions on relationships between education and cognition (Deary, Batty, Pattie & Gale, 2008; Calvin, Crang, Paterson & Deary, 2014; Paterson, Pattie & Deary, 2010, 2011).

No members of the SMS1947 had been contacted since youth, and their data had been largely unused for decades (Deary et al., 2009). The 6-Day Sample study has two ideas. First, it aims to re-contact surviving members of the 1,208-strong 6-Day Sample in their 70s, and to collect data on their lives since childhood and on their current health and cognitive capabilities. Second, it aims to link the childhood data from the whole Scottish Mental Survey 1947 (70,805-strong) anonymously to health records in the UK. We next go into more detail on each of these broad aims.

The first aspect of the study meant that the 1,208 members of the 6-Day Sample had to be traced through the NHS Central Register in Scotland, and in England and Wales. We considered that we could not contact people who had emigrated permanently from the UK. Then, identified, surviving members of the 6-

Day Sample would be invited, through the Registrar General for Scotland and their equivalent in England and Wales, to a follow-up study, carried out by post and telephone. This would involve answering questions about their lives, health and wellbeing, carrying out some basic physical tests at home, and completing some simple cognitive tests over the telephone. Sub-groups of the follow-up study participants would be invited to three sub-studies: an in-depth life course interview, a validation clinic visit, and a urinary cortisol study.

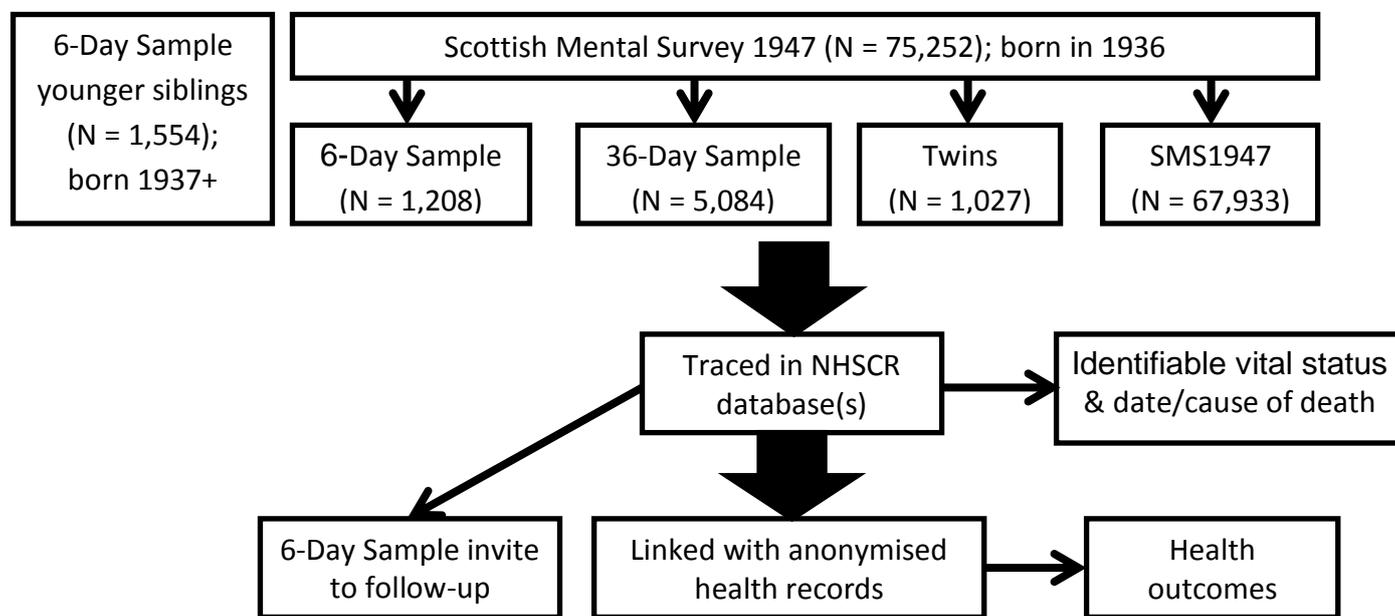
The second aspect of the study involved proposed linkage to UK health records: all members of the 6-Day Sample study would be located in the Scottish Morbidity Records and equivalent databases in England and Wales. Information within these databases would be linked anonymously (to the researchers) with the information gathered during childhood and early adulthood. Given the effort that this would require, it was decided to undertake this process not only for the 6-Day Sample but for the entire Scottish Mental Survey 1947 group (70,805 individuals, which increased to 75,252 when we included those who did not sit the intelligence test in 1947), and the 6-Day Sample's younger siblings.

Childhood intelligence data were available for the majority of these individuals and the study plan represents what could be the most comprehensive epidemiological study of the associations between

childhood intelligence and subsequent health outcomes ever conducted.

Figure 2 illustrates the study aims in diagrammatic form.

Figure 2. Diagram of the 6-Day Sample study aims



The permissions process: summary

A summary of the organisations we applied to is provided in Table 1 and a narrative of the process is in Appendix 1 (available in the Supplementary File). In total, 7 separate regulatory bodies or organisations were applied to for permission to receive data from four different sources: the NHS Central Register (NHSCR), Scottish Morbidity Records (SMR), Hospital Episode Statistics (HES) and the Office for National Statistics (ONS). Seven amendments or re-submissions were made as a result of changes to the study protocol or other supporting documents. The name and function of two of the bodies in England and Wales changed while our application was being processed. In total, 210 documents were sent in support of applications, amendments or re-submissions (not all of which are detailed in Table 1). From beginning to end, the process took 538 days between January 2012 and July 2013. Figure 3 shows a timeline of the permissions process.

The permissions process: ethical issues

The main ethical issues in our study were capacity and consent.

First, there was the possibility that, due to their age, some participants might be suffering from

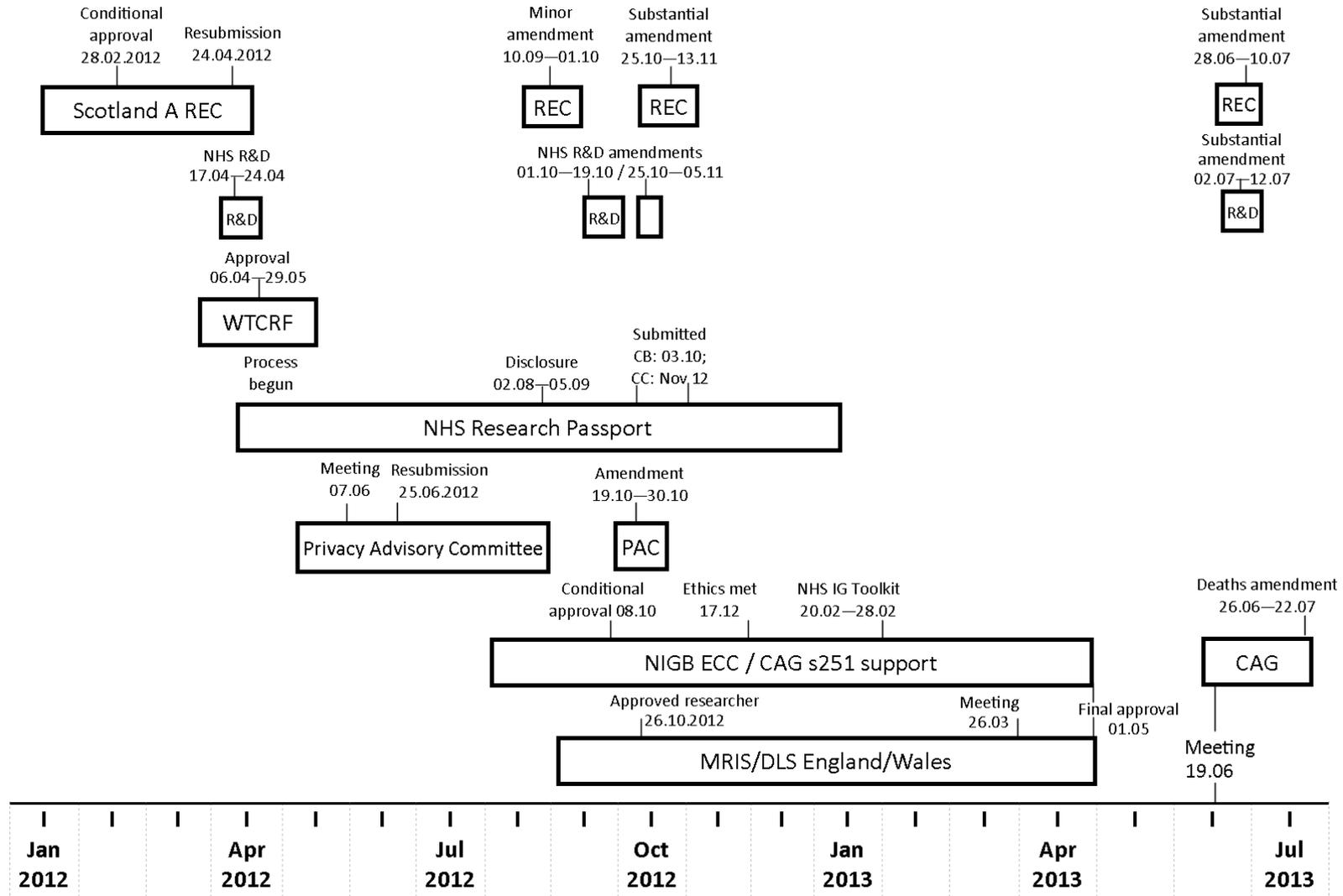
dementia or other disease that might impair their ability to give informed consent for the follow-up study. In order to maintain the representativeness of the 6-Day Sample, we wished to include these people in the study where possible, and we received ethical approval from a national Research Ethics Committee (REC) in Scotland to do so. However, participants in England/Wales were subject to different legislation and, rather than delay recruitment further by applying to an English REC, we elected to exclude the (very small) number of participants lacking mental capacity in England/Wales.

Second, the study relied on our using personal data without explicit consent from participants. For the linkage aspects, this was necessary in order to maintain the representativeness of the sample – not only would it be practically impossible to obtain consent from relatives of participants who had died, or from participants who had emigrated, but obtaining consent might have introduced systematic bias into our sample (Al-Shahi Salman et al., 2014). For the follow-up study, personal data held within NHSCR databases was required to send out the invitations – although we ourselves did not receive this information directly.

Table 1. Summary of the permissions process

Organisation	Permission sought	Supporting documents	Time to initial approval	Amendments	Number of documents for amendments	Time to final approval
NHS Research Ethics Committee Scotland A	General ethical approval	29	86 days	2	13	268 days
NHS Lothian R&D	General study approval	27	7 days (NB: REC approval first)	2	12	202 days
Wellcome Trust Clinical Research Facility	Follow-up study approval	13	53	0	N/A	53 days
NHS Research Passport	Permission to speak to participants	8	275 days	0	N/A	275 days
Privacy Advisory Committee	Data custodian for Scottish Morbidity Records & NHSCR data	30	113 days	1	2	169 days
Ethics and Confidentiality Committee of the NIGB (now the Confidentiality Advisory Group of the HRA)	Support under section 251 of the Health and Social Care Act (2006) to use health data without consent	49 (including NHS IG Toolkit)	203 days	1	0	339 days
Data Linkage Service, Health and Social Care Information Centre	Data custodian for Hospital Episode Statistics & ONS deaths and cancer data	26	230 days	1	11	230 days
Summary	--	182	--	7	38	538 days

Figure 3. Timeline of the permissions process for the 6-Day Sample study



Acronyms used: REC = Research Ethics Committee; R&D = NHS Lothian Research & Development; WTCRF = Wellcome Trust Clinical Research Facility; PAC = Privacy Advisory Committee; NIGB = National Information Governance Board; ECC = Ethics and Confidentiality Committee; MRIS = Medical Research Information Service; DLS = Data Linkage Service; CAG = Confidentiality Advisory Group.

In Scotland, permission was sought from the Privacy Advisory Committee, which provides advice, on a non-statutory footing, to the owners of Scottish health data, the NHS Information Services Division (ISD) and National Records of Scotland. In England/Wales, permission was sought from the Ethics and Confidentiality Committee of the National Information Governance Board (this has since been renamed and repositioned as the Confidentiality Advisory Group of the Health Research Authority). This body has statutory powers, under section 251 of the Health and Social Care Act 2006, to give approval for the use of medical data without consent.

In their review of current practices for the use of personal information for medical research in the UK, Laurie and Sethi (2011) summarised it as a 'consent or anonymise, with authorisation' approach. The 6-Day Sample study incorporated all three approaches – consent for the follow-up study, in which individuals provided data directly to us, anonymisation for the data linkage aspects, and authorisation for the provision of identifiable death data, for which neither consent nor anonymisation were appropriate.

The permissions process: challenges

A main challenge to us as researchers was gaining understanding of the legislation and framework for data linkage. This was not helped by the complexity of the current landscape, and the often unclear and conflicting guidelines (Laurie & Sethi, 2011). Although we embarked on a steep and successful learning curve, navigating the complexities inevitably delayed the process. One inconsistency which caused delay was the differing interpretation of the Statistics and Registration Service Act 2007 between Scotland and England/Wales. This led to some confusion over whether we were to receive identifiable or anonymised cause of death data.

Another substantial challenge was information governance. As researchers, we have a duty to ensure that the data provided by research participants is taken care of. To this end, practices were already in place to protect the data, store it securely, and prevent unauthorised access or loss. However, the majority of these practices were transmitted and maintained by members of the research team and had not been written in the form required by the external organisations. We

submitted our Ethics and Confidentiality Committee (ECC) application on Friday 17th August 2012, having spent several weeks producing a written 'System Level Security Plan' as required. On Monday 20th August 2012 the NHS Information Centre put into force a rule that stated that all researchers applying for secondary use of health data needed to complete the NHS Information Governance Toolkit. The full implications of this were not clear until many months later. The Toolkit is designed for completion by NHS Trusts in order to ensure that they are complying with legislation and the common law duty of confidentiality. We were advised to complete this at an institution level – i.e. on behalf of the whole of the University of Edinburgh. However, despite being part of a large institution with the associated benefits in terms of infrastructure and expertise, the existing University-wide data protection policies applied mostly to the handling of student data and were not sufficient for the handling of the particular research data of our project. In addition, the University has a diverse research community and practices and infrastructure varied between different research groups and departments. As a result, we were obliged to complete the Toolkit on behalf of just our research study. Although we only had to make minimal alterations to our existing practices, the Toolkit required evidence of all information governance practices. With some assistance from computing officers, we created 19 new documents in order to meet the Toolkit's requirements. And we passed.

The final challenge was perhaps the most intriguing. Although we weren't the first study to conduct data linkage on individuals in Scotland and in England/Wales, it transpired that we were the first study to request data from the three countries on the same individuals across time. Therefore, considerable effort was expended in fostering collaborative relationships between the organisations providing the data, and putting systems in place to ensure smooth linkage between the two external data sets and our own existing data.

Discussion

This case study—which is fully described in the narrative account in Appendix 1—outlines the lengthy process involved in obtaining permission to link life course data from an existing research study

with nationally-held health databases. Nevertheless, all permissions were eventually granted and the linkage and follow-up is now underway. This is not the first attempt in the UK to 'revive', many years later, a study first conducted in childhood. Both the Boyd-Orr cohort (Gunnell, Frankel, Nanchahal, Braddon & Davey Smith, 1996; Martin, Gunnell, Pemberton, Frankel & Davey Smith, 2005) and the Aberdeen Children of the 1950s study (Batty et al., 2005) accomplished this successfully. The complexities of tracing and contacting research participants decades later contrast with cross-sectional studies taking a snapshot of a population sample at a specific time-point, and those longitudinal cohort studies that have successfully maintained continuous contact with participants since baseline testing. As the second-oldest birth cohort in the UK with childhood data (the oldest being its predecessor, the Scottish Mental Survey 1932 (Deary et al., 2009; Scottish Council for Research in Education, 1933)), and one of only a handful of year-of-birth cohorts in the world, the Scottish Mental Survey 1947 offers a rare opportunity to link childhood intelligence (and, to a lesser extent, deprivation) with future health outcomes across the life course and into older age. The sub-groups of the SMS1947, in particular the 6-Day Sample, represent an opportunity to investigate the influences of psycho-social circumstances, social and environmental exposures, and occupational characteristics in early adulthood on subsequent health and social mobility.

We began this process with a novel yet simple idea: to trace an entire year-of-birth population and link information collected in childhood with subsequent health outcomes, and to invite a sub-group of this population to a follow-up study of health and wellbeing in old age. The process has been long with many challenges along the way. Protecting the privacy and confidentiality of personal data held within administrative databases is paramount. Most health data is routinely collected for the purposes of monitoring a population's health and evaluating health service use and utility. Although this is done without individuals' explicit consent, the common law duty of confidentiality means that individuals have the right to expect their personal data collected during the course of any contact with health services to be protected and their confidentiality maintained. Regulating access to health data is a necessary

means of achieving this. Whereas no researcher would disagree with the need for regulation, it could be argued that, at times, different interpretations and implementations of legislation governing personal health data adds unnecessary complexity to the process of obtaining approval.

Recommendations

Reflecting on our experience, there are a number of recommendations we could make that would streamline the permissions process and reduce wastage.

A data linkage information portal for researchers

Each country or administrative area could create an information centre as a single point of entry for researchers wishing to use their data for medical research. This portal should detail all the existing data sets, their ownership and provenance, access restrictions and information on how researchers can obtain permission to access the data. There is precedent for such a centre; the Finnish Information Centre for Register Research (<http://rekisteritutkimusen.wordpress.com/>) fulfils this purpose in Finland and linked administrative data sets such as that in Western Australia (<http://www.datalinkage-wa.org/>) also use this kind of entry point. However, these are the exceptions rather than the rule.

A centralised system for applying for permissions

Many countries have recognised and responded to the need for a more streamlined process of ethical review (Al-Shahi Salman et al., 2014); however, for the most part, the complexities of regulation and legislation involved in data linkage have to date prevented this from being extended to include review of applications to access health data for research purposes. An important aspect of this system would be the review of supporting documents, which frequently change in response to feedback from regulatory bodies and need to be re-reviewed by ethics committees. This process, for our study at least, was time-consuming and prone to human error – for example, simply forgetting to double check the version number and date on the 29th supporting document led to a two week delay in obtaining NHS R&D approval. A more efficient system would involve uploading supporting documents and submitting them for review centrally.

More guidance and training for researchers

Many researchers, like ourselves, are unfamiliar with the legislation and processes surrounding the use of data for which we are not directly responsible, and would benefit from outside guidance. Ideally, this would involve a single point of contact throughout the process with someone with expertise in the area. Although most researchers are already taking steps to protect their own research data, basic information governance training – including an overview of relevant legislation – should be mandatory for all researchers embarking on this process. Experienced researchers themselves might be a rich source of guidance and training. Initiatives such as Cohort and Longitudinal Studies Enhancement Resources (CLOSER: www.closerprogramme.co.uk) provide a platform for collaboration between studies, enabling them to pool resources and expertise, provide training, facilitate data linkage, encourage harmonisation across studies and increase the impact of longitudinal studies on policy and practice. The case study presented here has already provided input to the Scottish Government's data linkage strategy and plans (Brett & Deary, 2013; Deary, 2013).

A single repository of linked datasets which researchers can access in a secure environment

As with the existing linked administrative datasets in Western Australia, Finland and other countries, linking health and non-health datasets using a single identifier reduces the need for personal identifiers, reduces the likelihood of linkage errors, and protects the privacy of individuals. Accessing data in a secure environment or 'safe haven' reduces the need for such strict information governance requirements within a researcher's own environment, and enables access to those who lack the infrastructure to provide a secure environment. However, secure environments need to meet researchers' needs. Occasionally, several researchers within a team work in parallel on the same dataset and need to be able to pass results or even data between themselves. The data provided to researchers needs to be in a format they can use for analysis to avoid unnecessary and costly hours spent manipulating the data within the safe haven. Often, researchers wish to use their own written software or code in order to manipulate data, which can be prohibited or difficult to implement within a secure environment. The Secured Unified Research Environment (SURE), a remote-access computing facility funded by the Australian and New South Wales Governments as part of the Population Health

Research Network, provides a blueprint for a researcher-friendly secure environment (<https://www.sure.org.au>). Where a secure environment is not available or practical (for example, where researchers are geographically distant from a suitable safe haven), legally binding data use agreements provide an alternative. Unfortunately, these do not always ensure that the organisation holding the data retains control over its use, which can lead to difficulties such as that currently experienced by the HSCIC in England (House of Commons Health Committee, 2014b).

The management of public perception of the use of personal data for medical research

Public perception of the use of personal data, either with or without consent, is influenced by the media. Decisions on the use of personal data for medical research, especially without consent, rely on a delicate balance between the protection of privacy and confidentiality, and research carried out for the 'public good' (Gissler & Haukka, 2010; Laurie & Sethi, 2011; Regidor, 2004). All too often, this is inadequately portrayed in the media and focus lands on the negative aspects of data used in this way. In the UK, the Prime Minister David Cameron's enthusiasm for opening up NHS data has taken the form of the care.data project, which aims to link all NHS data together and greatly enhance medical research and permit access to the private sector. However, some argue that its implementation to date has left a lot to be desired, with the media focussing on the commercial exploitation of health data (Goldacre, 2014a, 2014b). The result is that the project has been 'put on hold' (House of Commons Health Committee, 2014a, 2014b) and all access to HES data is currently under review (Whitfield, 2014). In contrast, the Western Australian Data Linkage System (WADLS) has, almost from its inception, encouraged public involvement in research based on linked data through active consultation, public surveys, consumer representation on management committees and the publication of regular bulletins in the public domain outlining the WADLS and its research findings (Brook, Rosman & Holman, 2008; Holman et al., 2008).

Data linkage and data sharing have important roles to play in life course epidemiology. Linked data have the power to identify trends in population health and understand the aetiology of disease, and to inform better decisions about care and improve quality and safety standards within health and social care. Data linkage can unlock the potential of routinely-collected

health and non-health data dating back up to 50+ years. The increasing use of data linkage necessitates important decisions about ownership, access and confidentiality. Stringent controls need to be in place to ensure that access to health records is granted according to public interest and confidentiality is maintained throughout. However, as evidenced by the present case study, this process is often complex due to inconsistencies in legislation and interpretation of legislation across countries and data custodians. The process of obtaining permission is lengthy and often difficult to navigate, which inevitably leads to delays and wastage (Al-Shahi Salman et al., 2014). In the 6-Day Sample study, wastage occurred where a consideration of the costs and effort required to obtain additional permissions (for example, to link to health records held in Wales and Northern Ireland) was considered disproportionate to the potential gains. Nevertheless, it is important for us to state that all individuals within the regulatory bodies were helpful and expressed interest in the study and its aims.

Aside from these legislative and administrative complexities, research using routinely collected health records is not without its problems. There are two main sources of bias arising from the use of health records. First, the quality of health records needs to be considered. Aside from the inherent bias arising from the inclusion of only individuals who consult health services, health records rely on accurate completion by clinicians and hospital coders. Studies have shown cause of death certification to be problematic in terms of omissions and errors (Middleton et al., 2011). Indeed, the process is currently undergoing reform in England and Wales (Office for National Statistics, 2012). Second, linkage errors occur, particularly where probabilistic matching techniques are employed (Harron, Wade, Gilbert, Muller-Pebody & Goldstein, et al., 2014). Separating

linkage from analysis, as generally occurs in research using linkage to health records, means that researchers are unable to evaluate the impact such errors might have on their results (Harron, Wade, Muller-Pebody, Goldstein & Gilbert, 2012; Harron et al., 2014).

The landscape of data linkage is constantly changing. Concerns have been raised over the revision of the EU Data Protection Directive, which may result in even more stringent controls on the use of personal data without consent, rendering studies such as the 6-Day Sample less likely to take place (Mansell, 2013). The situation in the UK is promising. Politicians and decision-makers are recognising the added value of data linkage (Cabinet Office and Prime Minister's Office, 2011; Walker, 2011) and encouraging funding to be channelled in this direction. Four [Farr Institutes of Health Informatics Research](#) are in the process of being set up across the UK. In Scotland, the Government commissioned a Data Sharing and Linkage Service in collaboration with data custodians and researchers, and its proposed design – as well as the proposed implementation of the Farr Institutes – meets many of the recommendations detailed above (Data Sharing and Linkage Service, 2013).

Given the great encouragement to undertake data linkage, and the initiatives that are taking place throughout the world, we hope this case study will be valuable for those researchers who, like our team's investigators, mostly come from the social and medical sciences and who wish to conduct life course research on human health and wellbeing. We concede that the process might be more straightforward for experts in data linkage. We offer this report as a guide and warning to researchers in the field, and as a stimulus to the greater simplification and harmonisation of organisations, processes and legislation.

Acknowledgements

Caroline Brett and the 6-Day Sample study are supported by a grant from the cross-council Lifelong Health and Wellbeing Initiative (G1001401, Lifelong health and wellbeing of the 'Scotland in Miniature' cohort: the 6-Day Sample of the Scottish Mental Survey 1947). Ethical approval for the study was given by the Scotland A-Research Ethics Committee; ref: 12-SS-0024. The work was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the Biotechnology and Biological Sciences Research Council (BBSRC) and Medical Research Council (MRC) is gratefully acknowledged.

References

- Al-Shahi Salman, R., Beller, E., Kagan, J., Hemminki, E., Phillips, R. S., Savulesco, J., MacLeod, M., Wisely, J., & Chalmers, I. (2014). Increasing value and reducing waste in biomedical research regulation and management. *Lancet*, 383, 176-185. [http://dx.doi.org/10.1016/S0140-6736\(13\)62297-7](http://dx.doi.org/10.1016/S0140-6736(13)62297-7)
- Batty, G. D., Morton, S. M. B., Campbell, D., Clark, H., Davey Smith, G., Hall, M., Macintyre, S., & Leon, D. A. (2005). The Aberdeen Children of the 1950s cohort study: background, methods and follow-up information on a new resource for the study of life course and intergenerational influences on health. *Paediatric and Perinatal Epidemiology*, 18, 221-239. <http://dx.doi.org/10.1111/j.1365-3016.2004.00552.x>
- Batty, G. D., Wennerstad, K. M., Davey Smith, G., Gunnell, D., Deary, I. J., Tynelius, P., & Rasmussen, F. (2007). IQ in early adulthood and later cancer risk: Cohort study of one million Swedish men. *Annals of Oncology*, 18, 21-28. <http://dx.doi.org/10.1093/annonc/mdl473>
- Batty, G. D., Shipley, M. J., Mortensen, L. H., Boyle, S. H., Barefoot, J., Grønbaek, M., Gale, C. R., & Deary, I. J. (2008). IQ in late adolescence/early adulthood, risk factors in middle age and later all-cause mortality in men: the Vietnam Experience Study. *Journal of Epidemiology and Community Health*, 62, 522-531. <http://dx.doi.org/10.1136/jech.2007.064881>
- Batty, G. D., Wennerstad, K. M., Davey Smith, G., Gunnell, D., Deary, I. J., Tynelius, P., & Rasmussen, F. (2009). IQ in early adulthood and mortality by middle age. *Epidemiology*, 20, 100-109. <http://dx.doi.org/10.1097/EDE.0b013e31818ba076>
- Blane, D., Netuveli, G., & Stone, J. (2007). The development of life course epidemiology. *Review Epidemiology Sante Publique*, 55, 31-38. <http://dx.doi.org/10.1016/j.respe.2006.12.004>
- Booth, J. N., Tomporowski, P. D., Boyle, J. M. E., Ness, A. R., Joinson, C., Leary, S. D., & Reilly, J. J. (in press). Obesity impairs academic attainment in adolescence: findings from ALSPAC, a UK cohort. *International Journal of Obesity*.
- Boyle, P., Feijten, P., Feng, F., Hattersley, L., Huang, Z., Nolan, J. & Raab, G. (2009). Cohort Profile: The Scottish Longitudinal Study (SLS). *International Journal of Epidemiology* 38, 385-392. <http://dx.doi.org/10.1093/ije/dyn087>
- Brett, C. E., & Deary, I. J. (2013, October). Navigating record linkage in Scotland and England and Wales: reviving the 6-Day Sample study. Paper presented at CPD session, Public Health Directorate, Scottish Government, Edinburgh, UK.
- Brook, E. L., Rosman, D. L. & Holman, C. D. J. (2008). Public good through data linkage: measuring research outputs from the Western Australia Data Linkage System. *Australian and New Zealand Journal of Public Health*, 32, 19-23. <http://dx.doi.org/10.1111/j.1753-6405.2008.00160.x>
- Cabinet Office and Prime Minister's Office. (2011). PM speech on life sciences and opening up the NHS. Accessed on 18th July 2013 from <https://www.gov.uk/government/speeches/pm-speech-on-life-sciences-and-opening-up-the-nhs>.
- Calvin, C. M., Crang, J. A., Paterson, L., & Deary, I. J. (2014). Childhood evacuation during World War II and subsequent cognitive ability: The Scottish Mental Survey 1947. *Longitudinal and Life Course Studies*, 5, 227-244. <http://dx.doi.org/10.14301/llcs.v5i2.255>
- Calvin, C. M., Fernandes, C., Smith, P., Visscher, P. M., & Deary, I. J. (2010). Sex, intelligence and educational achievement in a national cohort of over 175,000 11-year-old schoolchildren in England. *Intelligence*, 38, 424-432. <http://dx.doi.org/10.1016/j.intell.2010.04.005>
- Cooper, R., Frank, J., Leyland, A., Hardy, R., Lawlor, D. A., Wareham, N. J., Dezaux, C., & Inskip, H. (2012). Using cohort studies in lifecourse epidemiology. *Public Health*, 126, 190-192. <http://dx.doi.org/10.1016/j.puhe.2011.12.002>
- Data Protection Act. (1998). Chapter 29.
- Data Sharing and Linkage Service. (2013). Analysis of responses to the Technical Consultation on the design of the Data Sharing and Linkage Service. Edinburgh: Scottish Government.
- Deary, I. J. (2013, October). Reusing historical data: the Scottish Mental Surveys of 1932 and 1947. Scottish Government Data Linkage Framework: Data Linkage Exemplars, Public Acceptability, Research Results and Pathfinder Projects (Keynote Lecture). Edinburgh, UK.
- Deary, I. J., Batty, G. D., Pattie, A. & Gale, C. R. (2008). More intelligent, more dependable children live longer. *Psychological Science*, 19, 874-880. <http://dx.doi.org/10.1111/j.1467-9280.2008.02171.x>
- Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, 35, 451-456. <http://dx.doi.org/10.1016/j.intell.2006.09.003>
- Deary, I. J., Whalley, L. J., & Starr, J. M. (2009). A lifetime of intelligence: follow-up studies of the Scottish Mental Surveys of 1932 and 1947. Washington, DC: American Psychological Association.
- Declaration of Helsinki (1964). Recommendations guiding medical doctors in biomedical research involving human subjects. Adopted by the 18th World Medical Assembly, Helsinki, Finland. *The New England Journal of Medicine*, 271, 473.

- Directive 95/46/EC. (1995). Director 95/46/EC of the European Parliament and of the Council of 24 processing of personal data and on the free movement of such data. Official Journal of the European Communities No L, 38, 31-55.
- Dunn, H. L. (1946). Record linkage. *American Journal of Public Health*, 36, 1412-1416. <http://dx.doi.org/10.2105/AJPH.36.12.1412>
- Flowers, J., & Ferguson, B. (2010). The future of health intelligence: Challenges and opportunities. *Public Health*, 124, 274-277. <http://dx.doi.org/10.1016/j.puhe.2010.03.006>
- Foresight Mental Capacity and Wellbeing Project. (2008). Final project report. London: The Government Office for Science.
- Gissler, M., & Haukka, J. (2004). Finnish health and social welfare registers in epidemiological research. *Norsk Epidemiologi*, 14, 113-120.
- Goldacre, B. (2014a). The NHS plan to share our medical data can save lives – but must be done right. *The Guardian*, 21st February 2014.
- Goldacre, B. (2014b). Care.data is in chaos. It breaks my heart. *The Guardian*, 28th February 2014.
- Goldring, S., & Newman, J. (2010). The ONS Longitudinal Study – a prestigious past and a bright future. *Population Trends*, 139, 4-10. <http://dx.doi.org/10.1057/pt.2010.2>
- Gunnell, D. J., Frankel, S., Nanchahal, K., Braddon, F. E. M., Davey Smith, G. (1996). Lifecourse exposure and later disease: a follow-up study based on a survey of family diet and health in pre-war Britain (1937-39). *Public Health*, 110, 85-94. [http://dx.doi.org/10.1016/S0033-3506\(96\)80052-7](http://dx.doi.org/10.1016/S0033-3506(96)80052-7)
- Harron, K., Wade, A., Muller-Pebody, B., Goldstein, H., & Gilbert, R. (2012). Opening the black box of record linkage. *Journal of Epidemiology and Community Health*, 66, 1198. <http://dx.doi.org/10.1136/jech-2012-201376>
- Harron, K., Wade, A., Gilbert, R., Muller-Pebody, B., & Goldstein, H. (2014). Evaluating bias due to data linkage error in electronic healthcare records. *BMC Medical Research Methodology*, 14:36. <http://dx.doi.org/10.1186/1471-2288-14-36>
- Holman, C. D. J., Bass, A. J., Rosman, D. L., Smith, M. B., Semmens, J. B., Glasson, E. J., Brook, E. L., Trutwein, B., Rouse, I. L., Watson, C. R., de Klerk, N. H., & Stanley, F. J. (2008). A decade of data linkage in Western Australia: strategic design, applications and benefits of the WA data linkage system. *Australian Health Review*, 32, 766-777. <http://dx.doi.org/10.1071/AH080766>
- House of Commons Health Committee. (2014a). Oral evidence: Care.data database, HC 1105 Tuesday 25 February 2014.
- House of Commons Health Committee. (2014b). Oral evidence: Handling of NHS Patient Data 1105 Tuesday 8 April 2014.
- Kho, M. E., Duffett, M., Willison, D. J., Cook, D. J., & Brouwers, M. C. (2009). Written informed consent and selection bias in observational studies using medical records: Systematic review. *British Medical Journal*, 338, b866. <http://dx.doi.org/10.1136/bmj.b866>
- Kuh, D., Ben-Shlomo, Y., Lynch, J., Hallqvist, J., & Power, C. (2003). Life course epidemiology. *Journal of Epidemiology and Community Health*, 57, 778-783. <http://dx.doi.org/10.1136/jech.57.10.778>
- Kuh, D., Cooper, R., Richards, M., Gale, C., von Zglinicki, T., & Guralnik, J. (2012). A life course approach to healthy ageing: The HALCyon programme. *Public Health*, 126, 193-195. <http://dx.doi.org/10.1016/j.puhe.2012.01.025>
- Kuh, D., Cooper, R., Hardy, R., Richards, M., & Ben-Shlomo, Y. (2014). A life course approach to healthy ageing. Oxford University Press.
- Laurie, G., & Sethi, N. (2011). Information governance of use of health-related data in medical research in Scotland: Current practices and future scenarios. SHIP Core Programme 2 Working Paper No 1. Edinburgh: SHIP.
- Lynch, J. & Davey Smith, G. (2005). A life course approach to chronic disease epidemiology. *Annual Review of Public Health*, 26, 1-35. <http://dx.doi.org/10.1146/annurev.publhealth.26.021304.144505>
- MacPherson, J. S. (1958). Eleven-year-olds grow up: A seven-year follow-up of a representative sample from the 1947 Scottish Mental Survey. London: University of London Press.
- Marmot, M., & Brunner, E. (2004). Cohort profile: The Whitehall II study. *International Journal of Epidemiology*, 34, 251-256. <http://dx.doi.org/10.1093/ije/dyh372>
- Mansell, P. (2013). EU data-protection amendments could undermine health research. *PharmaTimes online*, October 24 2013. Accessed 7th March 2014 from http://www.pharmatimes.com/article/13-10-24/EU_data-protection_amendments_could_undermine_health_research.aspx.
- Martin, R. M., Gunnell, D., Pemberton, J., Frankel, S. & Davey Smith, G. (2005). Cohort profile: The Boyd Orr cohort – an historical cohort study based on the 65 year follow-up of the Carnegie Survey of Diet and Health (1937-39). *International Journal of Epidemiology*, 34, 742-749. <http://dx.doi.org/10.1093/ije/dyi124>
- Mason, J. K., & Laurie, G. T. (2010). *Law and medical ethics*, 8th edition. Oxford: Oxford University Press.
- Maxwell, James. (1961). *The level and trend of national intelligence: The contribution of the Scottish Mental Surveys*. London: University of London Press.

- Maxwell, James. (1969). *Sixteen years on: A follow-up of the 1947 Scottish Survey*. London: University of London Press.
- Melton, L. S. (1997). The threat to medical-records research. *The New England Journal of Medicine*, 337, 1466-1470. <http://dx.doi.org/10.1056/NEJM199711133372012>
- Mental Survey Committee. (1949). *The trend of Scottish intelligence: A comparison of the 1947 & 1932 surveys of the intelligence of eleven-year-old pupils*. London: University of London Press.
- Mental Survey Committee. (1953). *Social implications of the 1947 Scottish Mental Survey*. London: University of London Press.
- Mental Survey Committee. (1958). *Educational and other aspects of the 1947 Scottish Mental Survey*. London: University of London Press.
- Middleton, D., Anderson, R., Billingsly, T., Bande Mangaliso Virgil, N., Wimberly, Y., & Lee, R. (2011). Death certification: issues and interventions. *Open Journal of Preventive Medicine*, 1, 167-170. <http://dx.doi.org/10.4236/ojpm.2011.13022>
- Office for National Statistics. (2012). *Death certification reform: A case study on the potential impact on mortality statistics, England and Wales*. Retrieved from <https://www.gov.uk/government/publications/death-certification-reforms-update-newsletter-issue-2>
- O'Reilly, D., Rosato, M., Catney, G., Johnston, F. and Brolly, M. (2012). Cohort description: The Northern Ireland Longitudinal Study (NILS), *International Journal of Epidemiology*, 41, 634-641. <http://dx.doi.org/10.1093/ije/dyq271>
- Paterson, L., Pattie, A., & Deary, I. (2010). Post-school education and social class destinations in Scotland in the 1950s. *Longitudinal and Life Course Studies*, 1(4), 371-393.
- Paterson, L., Pattie, A., & Deary, I. J. (2011). Social class, gender and secondary education in Scotland in the 1950s. *Oxford Review of Education*, 37(3), 383-401. <http://dx.doi.org/10.1080/03054985.2011.564847>
- Pearson, H. (2011). Study of a lifetime. *Nature*, 471, 20-24. <http://dx.doi.org/10.1038/471020a>
- Power, C., & Elliott, J. (2005). Cohort profile: 1958 British birth cohort (National Child Development Study). *International Journal of Epidemiology*, 35, 34-41. <http://dx.doi.org/10.1093/ije/dyi183>
- Regidor, E. (2004). The use of personal data from medical records and biological materials: ethical perspectives and the basis for legal restrictions in health research. *Social Science & Medicine*, 59, 1975-1984. <http://dx.doi.org/10.1016/j.socscimed.2004.02.032>
- Richter, M., & Blane, D. (2013). The life course: Challenges and opportunities for public health research. *International Journal of Public Health*, 58, 1-2. <http://dx.doi.org/10.1007/s00038-012-0436-z>
- Roos, L. L., Menec, V., & Currie, R.J. (2004). Policy analysis in an information-rich environment. *Social Science & Medicine*, 58, 2231-41. <http://dx.doi.org/10.1016/j.socscimed.2003.08.008>
- Seematter-Bagnoud, L., & Santos-Eggimann, B. (2006). Population-based cohorts of the 50s and over: A summary of worldwide previous and ongoing studies for research on health in ageing. *European Journal of Ageing*, 3, 41-59. <http://dx.doi.org/10.1007/s10433-006-0022-4>
- Scottish Council for Research in Education. (1933). *The intelligence of Scottish children: A national survey of an age-group*. London: University of London Press.
- Scottish Government. (2010). *Linking social care, housing & health data. Data linkage literature review 2010*. Edinburgh: Scottish Government.
- Stephens, A., Breeze, E., Banks, J., & Nazroo, J. (2012). Cohort profile: The English Longitudinal Study of Ageing. *International Journal of Epidemiology*, 42, 1640-1648. <http://dx.doi.org/10.1093/ije/dys168>
- Tanner, J. M. (1966). Galtonian Eugenics and the study of growth: The relation of body size, intelligence test score, and social circumstances in children and adults. *The Eugenics Review*, 58, 122-135.
- The Caldicott Committee (1997). *Report on the review of patient-identifiable information*. London: Department of Health.
- Udjus, L. G. (1964). *Anthropometric changes in Norwegian men in the twentieth century*. Oslo: Universitetsforlaget.
- Wadsworth, M., Kuh, D., Richards, M., & Hardy, R. (2006). Cohort profile: The 1946 National Birth Cohort (MRC National Survey of Health and Development). *International Journal of Epidemiology*, 35, 49-54. <http://dx.doi.org/10.1093/ije/dyi201>
- Walker, D. (2011). *Administrative data linkage: the key to improving public policy?* Accessed 18th July 2013 from <http://www.alliance4usefulevidence.org/administrative-data-linkage-the-key-to-improving-public-policy/>.
- Weiss, A., Gale, C. R., Batty, G. D., & Deary, I. J. (2009). Emotionally stable, intelligent men live longer: The Vietnam Experience Study Cohort. *Psychosomatic Medicine*, 71, 385-394. <http://dx.doi.org/10.1097/PSY.0b013e318198de78>
- Whitfield, L. (2014). HSCIC reviews HES access. *Ehealth Insider*, 17th April 2014. Retrieved from: <http://www.ehi.co.uk/news/ehi/9364/hscic-reviews-hes-access>

