Review of data releases by the NHS Information Centre

Sir Nick Partridge

17 June 2014

Why this report has been written

1. As a non-executive director of the Health and Social Care Information Centre, I was asked by the Board on 5th March 2014 to lead a review of data released by the NHS Information Centre (NHS IC)[1], one of our predecessor organisations. From April 2005 to March 2013, its role had been to collect and manage health records data, including sharing it with third parties under data sharing agreements which restricted its use.

2. It disappoints me to report that the review has discovered lapses in the strict arrangements that were supposed to be in place to ensure that people's personal data would never be used improperly.

3. These lapses occurred before the HSCIC came into being and so it might be said that they are not the HSCIC's fault. However, that is beside the point. The lapses are very much our responsibility to address. The HSCIC has a new Board and largely new senior executive team, but it inherited many of the NHS IC's staff and procedures. Crucially it inherited information-sharing agreements with universities, companies, charities and other organisations that were given access to data by the NHS IC[1]. Any lapses in the procedures for data released under the stewardship of the NHS IC during the eight years to 31st March 2013 may continue to have implications for the handling of the data today.

4. Although this is a report to the HSCIC Board, I have tried to make it intelligible to patients and service users. It is their data that we guard and it is their trust that we must earn. I want therefore to introduce at an early stage some explanation of material that my fellow Board members already know well.

5. The NHS IC was a public sector organisation that collected information about health and social care in England. The information included data about patients and service users originating from hospitals, GPs, local authorities and other service providers and commissioners. Among its many duties, the NHS IC had legal authority to disseminate the data it collected, while safeguarding individuals' confidentiality. Most of this data was published in aggregated and anonymised form, stripped of personal attributes that might permit the identification of any individual. The NHS IC published thousands of reports of this uncontroversial nature.

6. The NHS IC also had a statutory power to make data which might potentially identify individuals available for other organisations to analyse, with the requirements that they had a legal basis to do so and signed a legally binding information sharing agreement. The conditions in such agreements included restrictions on what information can and cannot be used for, how it must be stored securely and how it must eventually be destroyed.

7. The data provided to these other organisations under data sharing agreements is not anonymised. Although names and addresses are normally removed, it is possible that the identity of individuals may be deduced if the data is linked to other data. That is why restrictions on how the data is used must be rigorously observed.

8. I am aware that most people have little knowledge about what these statistical pieces of information are. They may even think that they include obviously personal

---

[1] Known as the NHS Information Centre (NHS IC), it was set up in 2005 as a Special Health Authority with a corporate identity of the HSCIC.

descriptions of either patient or illness. This is not the case and I have included examples of the type of data provided by the NHS IC in Appendix 1 of this report. This shows what they look like and makes clear that they consist of codes and numbers.

9. The current HSCIC Board is aware of the importance of keeping the flow of such data under close scrutiny. Minutes of the board meeting on 15th January 2014 recorded: "The Board requested a quarterly report of organisations that had requested or had contracts with the HSCIC for the provision of data. This would include the reason for the request, and the legal basis on which the data had been supplied."

10. This work gained added importance and urgency due to growing public interest in plans developed by NHS England to upload information from patients' GP records on to a national database to be run by the HSCIC, known as the care.data programme. The revelation that the NHS IC had released data to an actuarial body added to public concern, amidst a wide misunderstanding that personal medical records were being easily accessed by insurers. Further concerns expressed by the Health Select Committee on 25th February 2014 prompted the HSCIC Board to ask me to lead a review of all the data releases made by the NHS IC, its predecessor organisation.

What the review discovered

11. I commissioned Pricewaterhouse Coopers (PwC) to carry out independent inquiries into data sharing agreements between 1st April 2005, when the NHS IC was established, and 31st March 2013, when it went out of existence. I chaired a steering group which directed the work, ensuring a constant focus on bringing as much transparency as possible to a past that most people have little knowledge or understanding of.

12. The definition of "data release" used in this review was: "A release of data from which there is a potential risk of identifying people. This could be through direct identifiers or by providing enough information for self-identification or identification through combining with other sources. This would include a patient record level extract or a table of data with small numbers that have not been suppressed in line with an agreed standard." This means if there are small numbers of a rare disease in a postcode, it risks identifying individuals and so should not be published.

13. PwC focussed specifically on 3,059 releases of data that were contained in records maintained by the NHS IC. A breakdown of the types of organisation receiving this information is shown in Table 1.

Table 1:
Number of data releases to types of organisation, April 2005 – March 2013

879 data releases	to universities, typically for the purpose of research and analysis
827 data releases	to the Department of Health and the wider NHS, typically for the purpose of analytics, benchmarking and research
588 data releases	to a range of private sector organisations, typically for the purpose of analytics, benchmarking and research
358 data releases	to public corporations (e.g. Audit Commission), typically for the purpose of audit, analytics, benchmarking and research
125 data releases	to research organisations
84 data releases	to registered charities, typically for the purpose of research
50 data releases	to professional bodies (e.g. the Royal College of Surgeons in England), typically for the purpose of research
48 data releases	to government agencies, typically for a variety of research and analysis purposes
41 data releases	to central government departments for a variety of purposes, typically including research, analysis, census/population studies and benchmarking
33 data releases	to public bodies (e.g. Competition Commission), typically for the purpose of research and analysis
24 data releases	that were registered to an individual person in the Data Re-use Agreement or Data Sharing Agreement, instead of an organisation
2 data releases	where it was not possible to identify the organisation that received the data from information retained by the IC

Administrative failing

14. Perhaps the most surprising finding of this review is contained in the last line of Table 1. It shows two cases of data that was apparently released without a proper record remaining of which organisation received the data. Those examples of failed administration are not just of interest to archivists. Data of this type should not have been released without a data sharing agreement including restrictions on how the data should be stored, used and eventually destroyed - all of which should have been monitored by the NHS IC. It is impossible to monitor activity if there are not full, reliable records of the receiving organisations.

15. Staff are still trying to track down the two cases and believe that the lapses may have been harmless, in the sense that no identifiable or potentially identifiable data went missing. According to PwC, in one case the data "could have" gone from one part to another part of a company that was at the time contracted by the Department of Health to manage releases of Hospital Episode Statistics.

16. In the other case "it is likely that" data went to a nurse doing research work for a primary care trust in north-west England. Maybe it did. But this was not a system that was ever meant to rely on a maybe. To earn the public's trust in future, we

must be able to show that our controls are meticulous, fool-proof and solid as a rock.

17. It is a matter of fact that no individual ever complained that their confidentiality had been breached as a result of data being shared or lost by the NHS IC. The PwC inquiry found no evidence of malfeasance by any member of staff. We have learned of no attempt to get round the Data Protection Act. There was no complaint to the Information Commissioner's Office. Legitimate releases of data led to a large amount of good work by researchers, which has benefited medical science and contributed to improvements in the way health and social care services are managed. Yet that describes the system as it is meant to work. It does not excuse errors that, even if having no negative consequence, still would create concerns for the public about the controls that are in place.

18. I will turn now to the 3,057 releases of data which can be tracked to receiving organisations. PwC used a sampling technique to permit rapid analysis of the quality of governance applied to these releases. The technique used was standard PwC methodology, which it regards as professionally reliable. It is described more fully in Appendix 6 of the report.

19. The largest category of data released was from the Hospital Episode Statistics. An example of what a Hospital Episode Statistic looks like is included in Appendix 1 of this report. It consists of codes and numbers.

20. Between 1st April 2005 and 1st April 2009, all aspects of HES data release management had been outsourced to Northgate[2]. In the sample of releases analysed by PwC, all had some form of data agreement in place. However, PwC could not find evidence that Northgate got permission from the NHS IC before making releases, as it was supposed to do. PwC could not find confirmation that the appropriate approval committee authorisation was in place for any of the sample tested.

21. It is difficult to ascertain whether the problem was due to poor record keeping or procedural non-compliance. In the absence of evidence it would be unfair to conclude that Northgate was at fault, but it does appear that the NHS IC should have managed the contract with Northgate more tightly. Record keeping improved after the NHS IC took over management of data requests from 1st April 2009, but the PwC could not find records to confirm full compliance in about 10% of the sample.

22. Another large category of data releases went to support medical research, mainly in universities. Between 1st April 2008 and 31st March 2013 the NHS IC's Medical Research Information Service (MRIS) released data to 591 approved research projects. These releases were not one-off events. Typically researchers wanted to track a cohort of individuals over time to observe the progress of a medical condition or the effectiveness of a treatment. This was usually patient-identifiable information.

23. The size of the cohorts ranged from about 1,000 people up to 1.3 million. For some studies the tracking was done once a month, for others it was quarterly or done at longer intervals. The researchers specified the data they needed, which often

---

[2] The Department of Health (DH) awarded the contract to deliver the Hospital Episode Statistics (HES) service to Hays Redfern Limited (Hays) in December 2001. In August 2003, Hays plc sold Hays Redfern to Northgate Information Solutions, which became the contractor. When the IC came into being in April 2005 the contract was transferred from DH to the IC to manage.

included when a person died or other information collected and controlled by the Office for National Statistics (ONS). So, in addition to the normal approval procedures, some research programmes needed to go through a formal process known as ONS Legal Gateway approval.

24. Sample testing by PwC identified one research programme that had not gone through the gateway, but was still – in 2014 – being given access to ONS data. That was a sufficiently grave lapse for PwC to look into all the programmes that should have had ONS Legal Gateway approval.

25. It found a further eight that were continuing to be given access to mortality data without having gone through the gateway. The nine research studies were therefore unauthorised and have been suspended by MRIS while they seek gateway approval. This may take some while. In the meantime the researchers have not been told to delete data on advice from the ONS that this might be detrimental to many years of useful work. If they get ONS gateway approval, I understand that it is likely to be retrospective and so no long-term harm will have been done to medical research. However, the blunt truth is that important procedures were not followed. The lapse might not have been discovered if the HSCIC Board had not ordered this review. Since it has, there are consequences in the real world, with nine research programmes suspended, which are detailed in Appendix 2 of this report.

26. The full list of the 3,059 releases of data made by the NHS IC is attached to the PwC report. In reviewing this list, it was identified that there are four Data Sharing Agreements made by the NHS IC with three re-insurance companies which allow those re-insurers to continue to use the data until the agreements expire in 2015 and 2016. Only the original data was released and the HSCIC has not released any data to these companies.

27. Being aware of public concern about insurance companies holding data drawn from health sources and with the change in the law through the Care Act (2014), which restricts the flow of potentially identifiable data solely to purposes of benefit to the health and social care systems, I ensured that the HSCIC's Chief Executive wrote to the three companies concerned asking them to delete the data ahead of this legislation coming into force.

28. The PwC report covers 11 categories of data release and I will not go through all of them here. The conclusion that I draw from this evidence is that the system did not have the checks and balances needed to ensure that the appropriate authority was always in place before data was released. In many cases the decision making process was unclear and the records of decisions are incomplete. It also seems clear that the responsibilities of becoming a data controller, something that happens as soon as an organisation receives data under a data sharing agreement, were not always clear to those who received data. The importance of data controllers understanding their responsibilities remains vital to the protection of people's confidentiality.

29. There was no single gateway into the NHS IC where data requests could enter and there were too many disparate, disjointed processes for the sharing of data. The process for ensuring the appropriate deletion of data at the end of an agreement was inadequate. And, on a more technical point, there was confusion regarding the differences between a Data Sharing Agreement and a Data Re-use Agreement, and in what circumstances they were appropriate. The upshot of all these defects is that it is not clear if data has been released for appropriate purposes in all cases.

30. PwC did not criticise NHS IC staff for failing to appreciate the importance of robust information governance. Indeed PwC said the failings in terms of the processes, control and overarching governance were not systemic. However, the consequence of all the defects set out above is that it is not clear if data has been released for appropriate purposes in all cases. That is clearly unsatisfactory and along with the other issues of appropriate controls over data deletion and approvals, leads me to make nine recommendations for action. Before I set these out I want to explain the role of the National Back Office.

National Back Office

31. The National Back Office (NBO) is in Southport because that was where the original national population identity database was established by the Registrar General following the outbreak of war in 1939. An audit trail was maintained throughout the war years and when the NHS was established in 1947, the same information was used to allocate NHS numbers. The NHS Central Register has continued to supply NHS numbers since and was hosted by the Office of National Statistics until 2008 when it came under the control of NHS IC and was retitled the NBO. It transferred to the HSCIC in 2013.

32. The NBO provides a service for clinicians by identifying and linking each NHS patient in England, Wales and the Isle of Man to the care records uniquely associated with that person, and correcting confusions, duplications and inaccuracies. The NBO primarily helps clinicians to identify patients quickly, easily and accurately. When doctors are making important decisions about a John Smith, they can be confident they have the right John Smith's medical details in front of them.

33. That assistance to clinicians did not form part of PwC's review. However, a small but significant part of the NBO also responds to properly sanctioned tracing requests from law enforcement agencies such as the UK Border Agency (UKBA)/Home Office, the police, the former Serious Organised Crime Agency (SOCA) – now National Crime Agency (NCA), and the courts. When tracing information passes from the NBO to any of these authorities, it amounts to a form of data release and therefore fell within the ambit of the PwC review.

34. It should be stressed that the information never contains clinical information. The NBO did not, and does not, pass any clinical information to the police, UKBA or SOCA/NCA. If the NBO agreed to provide a trace in response to a lawfully sought request from one of these law enforcement agencies and if the trace was successful in identifying the individual concerned, the NBO offers only very limited information about the health area in which the person was, or was last, registered with a GP. If a law enforcement agency wanted to follow that up, it would have to approach the local health authority with a properly and lawfully obtained request (which during 2008-2013 would have been the Primary Care Trust.) In rare cases a law enforcement agency might be told that a person had died, but not given clinical information. The NBO may provide fuller information to the courts, including addresses and details of the individual's GP, if that is required to comply with a court order.

35. In the period 2008-2013 the law enforcement authorities made 28,744 trace requests. The NBO made 10,647 successful traces and there were 2,307 cases of the trace request being accepted, but no record of whether or not it was completed successfully.

36. Information was provided to the police and SOCA only if senior officers could demonstrate that it was needed to detect or prevent the most serious crimes. Information was typically provided to the UKBA to help with the detection or prevention of immigration offences, trafficking and abuse of the NHS. Over the period in question 59% of the successful traces for which records remain in existence were to assist the UKBA.

37. I was a member of Dame Fiona Caldicott's review of information governance and one of the core principles we established during this work was that there should be "no surprises" for the public in how information is used. They have a right to know and to be reassured that no clinical information is being disclosed. I therefore believe that the HSCIC should continue to publish the numbers of these enquiries received and responded to. This can be included in the quarterly update of the Register of all HSCIC data releases which first published on April 4 this year.

38. In the light of the above, I recommend to the HSCIC the following measures:

Recommendations

1) That the HSCIC undertakes a programme of work to ensure that data has been deleted appropriately for all data releases referenced in the PwC report, where the agreement has ended.

2) That the HSCIC develops one clear, simple, efficient and transparent process for the management of all data releases.

3) That the HSCIC implements a robust audit function, which will enable ongoing scrutiny of how data is being used, stored and deleted by those receiving it.

4) That the HSCIC publishes its policy, process and governance for the release of data.

5) That the HSCIC ensures there is clear, transparent and timely decision making, via the appropriate governance for all data releases, and that all decisions are documented and published on its website.

6) That the HSCIC implements a robust record keeping approach and that the details of all data releases (including the purpose for which they are released) are made available on its website.

7) That the HSCIC develops one Data Sharing Agreement, which is used for all releases of data, and which includes clear sanctions for any breaches.

8) That the HSCIC actively pursues a technical solution to allow access to data, without the need to release data out of the HSCIC to external organisations.

9) That the HSCIC quarterly Register of all data releases includes the number of law enforcement agencies' person tracing requests processed by the National Back Office. The Register will also include all data being released under NHS IC data sharing agreements, ensuring it is providing a comprehensive account to the public of all data being shared.

Appendix 1a: Example of cervical screening data provided as part of an existing data sharing agreement – NOT REAL DATA

KC53 PartA2 - Population status

PCT	Age	Resident Population	Number of women with recall ceased for:			Eligible Population
			Clinical reasons	Age reasons	Other reasons	
5AT	65-69	2832	1	0	0	2831
TAP	40-44	10623	543	0	65	10615

KC53 PartB - Women invited

PCT	Age	Women invited in the year as a result of:		Women invited in the year as a result of repeat in < 3 years for reasons of:		
		Call	Recall	Surveillance	Abnormality	Inadequate
5AN	70-74	0	8	1	1	0
5LN	50-54	110	1329	413	98	76

KC53 PartC2 - Women tested by result

PCT	Result of test	Number of women aged 25-64 tested in the year :		Repeat in < 3years for reasons of:			Number of women aged 25-64 tested in the year:		
		As a result of call	As a result of routine recall	Surveillance	Abnormality	Inadequate test	While recall suspended	While recall ceased	Not invited by programme
5A2	Severe dyskaryosis ?invasive	0	2	0	1	0	0	0	0
TAM	Negative	963	6823	1392	283	92	103	45	2921

KC53 PartD - Result of test by age

PCT	Age	Woman's most severe test result in the year						
		Negative	Borderline	Mild dyskaryosis	Moderate dyskaryosis	Severe dyskaryosis	Severe/ ?invasive	?Glandular neoplasia
5QW	70-74	24	0	2	2	1	0	0
5C1	45-49	4732	96	20	12	18	3	1

Appendix 1b: Fabricated HES sample, inpatient – NOT REAL DATA

ACTIVAGE	ADMIAGE	ADMIDATE	ADMIMETH	ADMINCAT	ADMINCATST	ADMISORC	ADMISTAT	ANAGEST	ANASDATE	ANTEDUR	BEDYEAR	BIRESUS_1	BIRESUS_2
34	33	1972-03-12	28	01	02	48	NULL	NULL	NULL	NULL	5	NULL	NULL
45	20	1982-03-12	11	02	99	66	NULL	NULL	NULL	NULL	NULL	NULL	NULL
7002	7002	1977-03-12	13	99	01	19	NULL	NULL	NULL	NULL	254	NULL	NULL

BIRESUS_3	BIRESUS_4	BIRESUS_5	BIRESUS_6	BIRESUS_7	BIRESUS_8	BIRESUS_9	BIRORDR_1	BIRORDR_2	BIRORDR_3	BIRORDR_4	BIRORDR_5	BIRORDR_6	BIRORDR_7
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL

BIRORDR_8	BIRORDR_9	BIRSTAT_1	BIRSTAT_2	BIRSTAT_3	BIRSTAT_4	BIRSTAT_5	BIRSTAT_6	BIRSTAT_7	BIRSTAT_8	BIRSTAT_9	BIRWEIT_1	BIRWEIT_2	BIRWEIT_3
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL

BIRWEIT_4	BIRWEIT_5	BIRWEIT_6	BIRWEIT_7	BIRWEIT_8	BIRWEIT_9	CARERSI	CATEGORY	CAUSE	CDSVERSION	CENDUR	CENSAGE	CENSTAT	CENWARD
NULL	NULL	NULL	NULL	NULL	NULL	01	10	F950	NULL	NULL	66	NULL	NULL
NULL	NULL	NULL	NULL	NULL	NULL	02	21	N994	NULL	1123	NULL	3	NULL
NULL	NULL	NULL	NULL	NULL	NULL	99	33	D101	NULL	NULL	NULL	NULL	NULL

CLASSPAT	CSNUM	CURRWARD	DELCHANG	DELINTEN	DELMETH_1	DELMETH_2	DELMETH_3	DELMETH_4	DELMETH_5	DELMETH_6	DELMETH_7	DELMETH_8	DELMETH_9
1	NULL	AB	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
2	NULL	CD	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
3	NULL	EF	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL

Appendix 1c: NCMP dummy data – NOT REAL DATA

Pupil_ID	Local_ID	Exclude flag	Sex	Age	School yr	Height	Height_z	Height_p	Weight	Weight_z	Weight_p	BMI	BMI_z	BMI_p
10492098	5ET-03457	0	F	65	R	104.5	-1.56815291	0.05842273	16	-1.41871009	0.07799178	14.6517	-0.57436853	0.28285922

BMI_p	BMI_class _pop_mon	BMI_class_ clinical	Mont_ meas	DayofMeasu rement	PCT	SHA	Schl_ GOR	Schl_Supergr oup_ONS_AC	Schl_Group _ONS_AC	Schl_Subgro up_ONS_AC	Schl_U rban	Schl_IMD_ Decile	Child_GOR
0.28285922	7	7	4	2	5ET	Q33	E	4	4.3	4.3a	6	7	E

Field Name	Field Description
Pupil_ID	Unique ID code for each pupil
Local_ID	PCT ID code for each pupil
Exclude_flag	Pupils attending independent and special schools are flagged as 1 (these records need to be excluded to match the published figures)
Sex	Sex of pupil
Age	Age of pupil (in months)
School_yr	School Year of pupil - derived from child age (R: Reception, 6: Year 6)
Height	Height of pupil (in cm)
Height_z	Height z score - derived from British 1990 growth reference, using Age, Sex and Height fields
Height_p	Height centile - derived from British 1990 growth reference, using Age, Sex and Height fields
Weight	Weight of pupil (in kg)
Weight_z	Weight z score - derived from British 1990 growth reference, using Age, Sex and Weight fields
Weight_p	Weight centile - derived from British 1990 growth reference, using Age, Sex and Weight fields
BMI	BMI of pupil in kg/m2 - derived from height and weight
BMI_z	BMI z score - derived from British 1990 growth reference, using Age, Sex and BMI fields
BMI_p	BMI centile - derived from British 1990 growth reference, using Age, Sex and BMI fields
BMI_class_pop_mon	BMI classification to UK90 population monitoring centiles (85th/95th centiles for overweight and obese), plus 2nd centile for underweight (See table BMI_class_pop_mon)
BMI_class_clinical	BMI classification to UK90 clinical centiles (approx 91st/98th centiles for overweight and obese), plus 2nd centile for underweight (See table BMI_class_clinical)
Month_meas	Month of measurement (1: January, 2: February etc)
DayofMeasurement	Day in week measurement took place (1: Sunday, 2: Monday, 3: Tuesday, 4: Wednesday, 5: Thursday, 6: Friday, 7: Saturday)

PCT	PCT that submitted measurement (see table PCT_ data) based on PCT codes as at 1st April 2011
SHA	SHA - based on PCT that submitted measurement
Schl_GOR	Government Office Region - derived from postcode of school (see table GOR_codes)
Schl_Supergroup_ONS_AC	ONS Area Classification Supergroup code - derived from school postcode (see table ONS_AC_Supergroup)
Schl_Group_ONS_AC	ONS Area Classification Group code - derived from school postcode (see table ONS_AC_Group)
Schl_Subgroup_ONS_AC	ONS Area Classification Subgroup code - derived from school postcode (see table ONS_AC_Subgroup)
Schl_Urban	ONS Urban/ Rural classification - derived from school postcode (see table Urban_rural)
Schl_IMD_decile	The decile that the Index of Multiple Deprivation (IMD) 2010 score falls into - derived from LSOA of school (0: not available, 1: 1st decile (least deprived),..., 10: 10th decile (most deprived))
Child_GOR	Government Office Region - derived from child postcode (see table GOR_codes)
Child_Supergroup_ONS_AC	ONS Area Classification Supergroup code - derived from child postcode (see table ONS_AC_Supergroup)
Child_Group_ONS_AC	ONS Area Classification Group code - derived from child postcode (see table ONS_AC_Group)
Child_Subgroup_ONS_AC	ONS Area Classification Subgroup code - derived from child postcode (see table ONS_AC_Subgroup)
Child_urban	ONS Urban/ Rural classification - derived from child postcode (see table Urban_rural)
Child_IMD_decile	The decile that the Index of Multiple Deprivation (IMD) 2010 score falls into - derived from LSOA of child (0: not available, 1: 1st decile (least deprived),..., 10: 10th decile (most deprived))
School_child_distance	The straight line distance between the school postcode and child postcode (in km).
Removal_criteria	Removal criteria (see table removal_criteria)

Appendix 2: MRIS - suspended longitudinal research studies

MRIS, the Medical Research Information Service, provides a service to researchers undertaking longitudinal studies. MRIS Of the live MRIS agreements, nine have been suspended as the review identified that they did not have the correct, up-to-date legal approvals in place. MRIS, the Medical Research Information Service, provides a service to researchers undertaking longitudinal studies. MRIS helps organisations, typically universities, to track cohorts of patients which typically range from around 1,000, with the biggest one being 1.3million.The suspended studies are historical studies that were originally fully covered by section 60 of the Health and Social Care Act 2001, but their approvals were not updated following a legislative shift in 2006. The studies are:

1.   Institute of Child Health

This study is to track and report on patients who have received or who are currently receiving Growth Hormone Therapy in order to assess the treatment.

2.   University of Cambridge: European prospective Investigation into Cancer.

This study investigates the link between food consumption and cancer.

3.   University of Oxford: Oxford Register of Early Childhood Impairments.

This study will use the information for planning services and support for children with disabilities such as cerebral palsy, deafness and blindness.

4.   Imperial College healthcare NHS Trust

This study involves longitudinal mortality data tracking of patients undergoing cardiothoracic surgical procedures at Hammersmith Hospital.

5.   University of Oxford

This study involved long term follow-up of breast cancer patients to examine the factors affecting survival of breast cancer.

6.   University of Birmingham

This study is to investigate the risk and benefits for women with early diagnosis breast cancer taking Adjuvant Tamoxifen for at least 3 years.

7.   University of Birmingham

This study involves the assessment of efficacy and safety of Donepezil and Aspirin in Alzheimer's disease.

8.   University of Cambridge

 This study is to create accurate information of the results of using different types of hydrocephalus shunts in the UK.

9.   Institute for Cancer Research

This study is to evaluate the effectiveness of the NHS Breast Screening Programme for England and Wales.