



coordinator@medconfidential.org

(Interim) Information Commissioner
Information Commissioner's Office
Wycliffe House, Water Lane
Wilmslow, Cheshire SK9 5AF

7th July 2016

Re: Complaint about objections to dissemination of data from the HSCIC

Dear Information Commissioner,

On 11th June 2015, we wrote to the Commissioner with a complaint about the dissemination of data from the HSCIC, with specific regard to what are known as 'Type 2' objections and the Hospital Episode Statistics (HES). On 23rd April 2016, HSCIC announced some changes as a result of an Undertaking with the Commissioner, and an associated Direction from the Secretary of State.

We are writing now to follow up on our original complaint given the way in which these changes have been implemented in practice, which HSCIC has confirmed to us in writing.¹ Our concern, and the concern of patients who continue to contact us about their opt outs, is that this implementation falls far short of what both the public and your Office would reasonably expect, and that data subjects' personal data continues to be processed unfairly.

There are three main issues:

1. Patients who have opted out are still having their data disseminated by HSCIC for purposes beyond their direct care

At the beginning of 2014, very few patients were aware that their data was being sold or shared beyond the NHS. As patients became aware of this during the course of the care.data debacle and further revelations about the practices of HSCIC and its precursor bodies, many chose to exercise their right to opt out.

As of April 2016, HSCIC confirmed that around 2.2% of patients in England² – roughly 1 in 45 patients – had objected to their data leaving HSCIC for any purpose beyond their direct care.³

¹ See letter from Prof Martin Martin Severs to medConfidential, dated 1 June 2016 - enclosure number 4.

² <http://www.hscic.gov.uk/catalogue/PUB20527/exp-care-info-choi-eng-ccg-apr-2016.pdf>

³ Many also objected to their data leaving their GP practice, but such Type 1 objections are outwith this complaint.

During the period January 2014 - April 2016, HSCIC continued to disseminate significant quantities of individual-level linked event data, including the data of patients who had made a Type 2 objection, in much the same way as it had been doing in the years before 2014. This data was disseminated (mostly as part of HES) for both commercial use and for commercial re-use, e.g. by 'information intermediaries' servicing both NHS and private sector customers – examples of the very practices that had caused public and Parliamentary outcry in 2014.

While it is understood that HSCIC could not access or act upon the Type 2 objections that had been registered on GP systems until Directed to do so in April 2016, once HSCIC could access the Type 2 objections the most reasonable expectation of patients who opted out is that their objections would be properly honoured and that their data would cease to be sold or shared by HSCIC for any purpose beyond their direct care.

However, it is now clear that the data of individuals who registered a Type 2 objection is *and will continue to be* disseminated by HSCIC as part of the Hospital Episode Statistics – which are not statistics in any normal sense of the word, but rather linked raw event data.⁴ This dissemination continues to include commercial use and commercial re-use, outside the direct control of the HSCIC, which is clearly contrary to the express wishes of patients who have opted out, and contrary to the public statements of the Secretary of State when offering the opt out to patients.

There is no technical or practical reason why HSCIC cannot remove the Type 2 objectors' data from HES releases, and re-release past data sets. It is the process we understood was to have happened back in April, before the reduction in scope emerged. And HES is the principle dataset of public concern – a fact of which HSCIC is completely aware.

Simply put, for those patients concerned enough to have opted out, whose instructions were then ignored for two years, very little meaningful will have changed. One of the primary mechanisms by which their data was previously being sold and shared (i.e. HES, which engaged demonstrable concerns when it came to public attention in 2014) will continue to operate, and their data will still be included – going indeed to some of the very same companies and organisations, under commercial use and commercial re-use contracts, for purposes beyond their direct care.

This being the case, we believe that these patients' personal data is and will continue to be processed unfairly, outside of their reasonable expectations.⁵

This is also not about the content of any future arrangements that may be derived after a future consultation. This is about the status quo, in operation, as we send this letter.

⁴ As clearly illustrated in Annexes to the Partridge Review, June 2014, e.g. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/367788/Sir_Nick_Partridge_s_summary_of_the_review.pdf

⁵ With regard to item (3) of the Undertaking: will HSCIC directly inform every patient who has registered a Type 2 objection that their individual-level data will continue to be included in HES, exactly as it was before they objected, or will it avoid explicit mention of HES and its ongoing dissemination?

2. Deeming over 25 years of linked hospital records “out of scope” of a patient’s expression of dissent will exacerbate the concerns that the Undertaking was designed to assuage, and undermine the authority of such Undertakings.

It is bad enough that, *despite having the capacity to do so*, HSCIC continues to disseminate Hospital Episode Statistics without removing the data of those who have registered a Type 2 objection. However, in ruling HES data “out of scope” for Type 2 objections, HSCIC will compound years of unfair processing with a policy and approach that evades key elements of the ICO’s Undertaking, and in so doing, undermines the regulatory regime the Commissioner’s Office is tasked with protecting.

To be absolutely clear, HSCIC’s stated position is as follows:

“The HSCIC policy position is that type 2 opt-outs do not apply where direct identifiers in the data sets have been removed or replaced with pseudonyms; and the data dissemination application has been approved for release through the end to end Data Access Request Service (DARS).”⁶

i.e. the position is clearly intended to apply to the dissemination of HES data.

We consider the detail of this policy and approach in more depth in part (3) of this complaint, but for the purpose of this part (2) emphasise that it was HSCIC’s and its precursor bodies’ sale or sharing of HES with commercial interests that informed many patients decision to opt out. This was perceived as an actual breach of trust by many, including Parliament, as opposed to the potential breach represented by proposals to extract patients’ GP data under the care.data programme.

We note that the focus of Sir Nick Partridge’s 2014 Data Release Review included HES data, to the extent that an annex included examples of HES data in order to raise public awareness of its content.⁷ HSCIC was, and is, clearly aware that HES is of public concern.

It is definitely the case that the operation of the Type 2 opt out had to be changed – due to the incorrect literal definition of the 9Nu4 code – as, despite what the public were told, it did not exclude the use of data for patients’ direct care. That is uncontentious. It is however deeply contentious, and an entirely separate decision, to change the application of the opt out so that HES is now deemed out of scope, when the most reasonable expectation from what people were told continues to be that it is in scope.

To deem HES out of scope of Type 2 objections, in effect ignoring or overriding patients’ reasonable expectation that their data will be excluded from HES, suggests HSCIC or the

⁶ Letter from Prof Martin Martin Severs to medConfidential, dated 1 June 2016 - enclosure number 4.

⁷ https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/367791/HSCIC_Data_Release_Review_PwC_Final_Report.pdf

decision making body has learned very little. It will frustrate many, whose trust has already been abused, and further undermine confidence in the NHS's handling of patient data.

In addition, if applied consistently, this policy and approach would appear to evade key elements of the ICO Undertaking – specifically, items (4), (5) and (6). If HES is deemed out of scope for Type 2 objections then HSCIC must believe it need not notify recipients of HES data that the datasets they hold may include records relating to patients who have chosen to opt out; that there is no requirement that any such data should not be disseminated further; and that recipients of such data have no need to destroy and/or replace the HES datasets they hold with new datasets with Type 2 objectors' data removed.

After deeming HES data out of scope, HSCIC may argue it is compliant with the terms of the ICO's Undertaking. It may arguably comply with the letter of what was written, while completely ignoring the spirit and intention: delivering on promises made to the public.

We believe this sets dangerous precedents, not only for the ongoing dissemination of the data of patients who have registered an objection, but also undermining the impact of the Undertaking and thereby the authority of the Commissioner's Office. Is it appropriate that a body can simply deem a significant proportion of its activities – some of the very activities which put it in breach in the first place – as out of scope of the only mechanism being offered to patients to register dissent?

3. HSCIC has failed to satisfy required standards

This third issue involves a more detailed discussion which we have broken into sections in an Annex below, but boils down to HSCIC's assertion that "Anonymised in Context" is a justification for continuing to share data – the data in question being the lifetime linked medical hospital history of an individual, including event dates, diagnoses and treatments – which HSCIC claims is compliant with the ICO's Anonymisation Code of Practice when protected by only a contract, with replacement of some direct identifiers with pseudonyms, but no modification of other data even HSCIC's own PIA describes as "high risk".

One of the fundamental underlying questions of this complaint is whether identifiable data can be considered "Anonymised" simply by issuing a legal agreement that this is the case, without full and proper consideration of the data that is subject to that agreement.

HES data is widely disseminated for a range of uses; this is the basis of public concern. It is also the basis for a great deal of political and commercial lobbying for data flows to continue just as they have for 25 years or more, regardless of fair processing or consent status. This pressure may explain why the scope of the opt outs appears to have been silently changed.

Quite simply, commercial entities with permission to re-use the data – through over-broad agreements and creative interpretation of language carefully chosen to leave loopholes, e.g. "solely commercial" and "for the promotion of health" – wish to continue making money off the data.

Now, regardless of these creative interpretations, HSCIC appears firm in its belief that it can send HES data to anyone who holds a contract, ignoring unambiguous patient dissent.

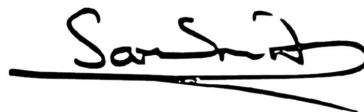
Given it is an administrative census of over two decades-worth of medical treatments in UK hospitals, is the Commissioner content that HSCIC's implementation and handling of HES is sufficiently "anonymised" according the intent and meaning of the Anonymisation Code of Practice, and other standards?

We continue to be happy to engage with you and your staff constructively on issues related to medical privacy, recognising, as Dame Fiona Caldicott said introducing her recent review, that there is "a spectrum of opinion, between those people who are very worried about confidentiality and their privacy, and those people who are very willing for information to be shared, both for themselves, but also for their families and wider community... The public is increasingly interested in what is happening with their information."⁸

Yours sincerely,



Phil Booth, medConfidential



Sam Smith, medConfidential

Cc: National Data Guardian,
HSCIC Caldicott Guardian,
Department of Health,
UK Anonymisation Network,
National Statistician/UKSA.

Enc: Copies of letters between medConfidential and HSCIC, in sequence:

- 1-medconf-to-hscic.pdf
- 2-hscic-to-medconfidential.pdf
- 3-medconf-to-hscic-caldicott-guardian.pdf
- 4-hscic-caldicott-guardian-to-medconfidential.pdf

⁸ <https://www.youtube.com/watch?v=iyfrjmetJrs>

Annex

[Background](#)

[1 billion records in the “Hospital Episode Statistics”](#)

[Dated information is identifiable data](#)

[Pseudonymisation is not anonymisation](#)

[Risk of re-identification](#)

[The ICO’s Anonymisation Code of Practice](#)

[The legacy of care.data: ignoring alternatives](#)

[What the public was told](#)

[Summary](#)

Background

For the sake of clarity, though we make occasional reference to Type 1 objections below, we reiterate that the issue here is not about GP-held data; the focus of this complaint is about the patient data from hospitals contained within the Hospital Episode Statistics. GPs are only tangentially involved, as they were the point of contact for patients to register dissent from having their data leave HSCIC for purposes beyond their direct care.

1 billion records in the “Hospital Episode Statistics”

According to HSCIC: *“Hospital Episodes Statistics (HES) is a data warehouse containing records of all patients admitted to NHS hospitals in England. It contains details of inpatient care, outpatient appointments and A&E attendance records.”*⁹ The HES Privacy Impact Assessment, published in May 2016, states: *“Many millions of records are stored, each one containing patient identifying details (including NHS Number) and confidential details about the patient’s care in hospital.”*¹⁰

And elsewhere, HSCIC says: *“Hospital Episode Statistics (HES) contains around 1 billion records on patients attending Accident and Emergency units, being admitted for treatment or attending outpatient clinics at NHS hospitals in England.”*¹¹ This figure of 1 billion records is supported in promotional material by commercial companies who continue to receive access to HES, and dates back to 2014.¹² One commercial reuse licensee boasts “we have access to 1.5 billion pseudonymised hospital episode statistic records dating back to 2006”.¹³ Previously, not all hospitals used the NHS Number as a consistent identifier. Following the Health and Social Care (Safety and Quality) Act 2015,¹⁴ all UK hospitals are now required to

⁹ <http://www.hscic.gov.uk/searchcatalogue?productid=20809&q=title%3a%22hospital+episode+statistics%22&sort=Relevance&size=10&page=1#top>

¹⁰ <http://www.hscic.gov.uk/article/7116/Consultation-on-the-Hospital-Episode-Statistics-Privacy-Impact-Assessment-Report>

¹¹ <http://www.hscic.gov.uk/hesdid> - as of 22 June 2016.

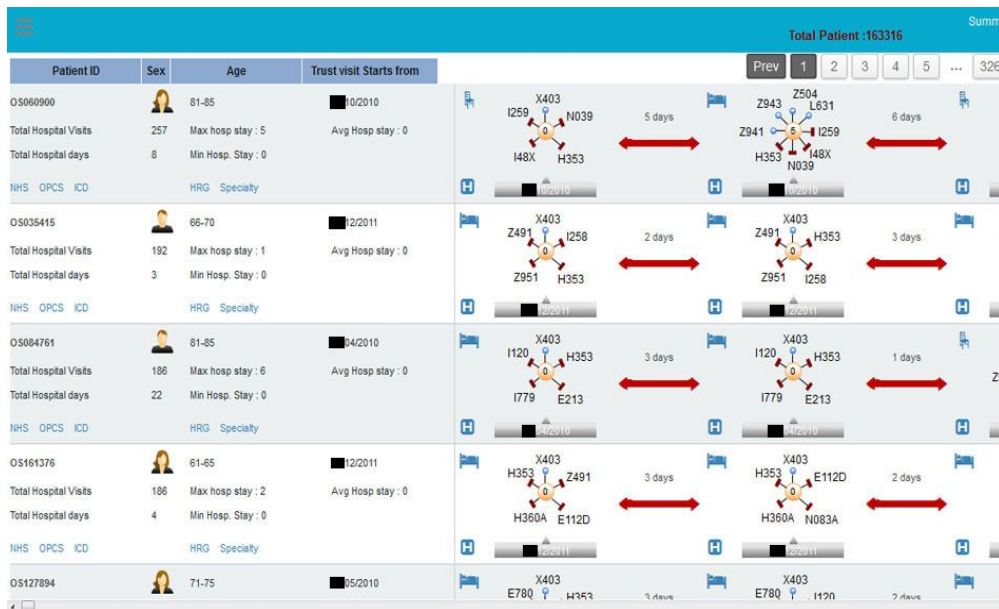
¹² e.g. https://medconfidential.org/wp-content/uploads/2014/04/one_billion.jpg

¹³ <http://www.harveywalsh.co.uk/what-we-do.aspx>

¹⁴ http://www.legislation.gov.uk/ukpga/2015/28/pdfs/ukpga_20150028_en.pdf

use the NHS number as a consistent identifier,¹⁵ leading to all hospital data being included in HES and to all patients being linkable across all hospitals.

The data can be reconstructed as below, from one of the commercial users¹⁶ of HES:



Each data item includes the date of the event, the coding of the type of the event (and optional value) and a patient identifier, a pseudonym, that links one patient's treatments to all other treatments for that patient.

HSCIC's current Data Release Register shows that commercial users and re-users continue to receive HES data in the same way as they did prior to the 2014 changes. The data of patients who have opted out continues to flow to these organisations.

Dated information is identifiable data

The first Caldicott Review, which continues to be applicable, provides a list of data items that are considered identifiable by the NHS. Annex 7 of the Review states (emphasis added):¹⁷

"The Working Groups identified a number of items by which a person's identity may be established. These include:-

- ...
- *Date of Birth*

¹⁵ <http://systems.hscic.gov.uk/infogov/iga/resources/considentifier.pdf>

¹⁶ The full dates of events have been redacted by medConfidential. Image at full size: <https://medconfidential.org/wp-content/uploads/2014/03/OmegaSolver.jpg>

¹⁷ http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4068404.pdf

- **Other Dates (i.e death, diagnosis)**
- **Sex...**

This being the case, it follows that the information contained within a single person's Hospital Episode Statistics entries must be considered identifiable for that person, even were the direct identifiers to be removed, and even if the data were not linked across their lifetime – which, for HES, it continues to be.

It is also relevant to note that with a longitudinal administrative census of health data, such as HES, it is known that if a patient was treated in a UK hospital, they will be in the dataset. This is not a survey with responses determined by the patient - it is an administrative census based on data collected for other purposes.

While “date of birth” may be admitted to have higher risk of re-identification, that single field relates only to the date of birth of the child. Entirely ignored from consideration, is that the mother who gave birth also has a number of dated medical events that appear in the various Hospital Episode Statistics datasets. Yet those other fields are not considered at all, despite being identical to the date of birth of the child. As such, should the birth dates of a woman's children would act as an almost certainly unique key into the entire hospital event history of the mother, linked through the pseudonym.

Pseudonymisation is not anonymisation

Paragraph 26 of the forthcoming General Data Protection Regulation states: *“Personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person.”* Were the Regulation now in force, the position would be entirely clear.

But the basic point is conceded, as indicated for example by recommendation (v) of HSCIC's Pseudonymisation Review interim report, which states (emphasis added): *“Where others access data held by the HSCIC, pseudonymisation should be applied in all appropriate circumstances, **alongside other techniques to minimise risk of re-identification.**”*

HSCIC's Data Release Register shows HES data relating to patient hospital treatments is released with no other techniques applied, i.e. the information provided is pseudonymised but otherwise unmodified administrative data. It may be that HSCIC has implemented what the Code describes as “Limited Access Safeguards”, i.e. some contractual constraints, but that alone is insufficient.

The NHS's own Anonymisation Standard for Publishing Health and Social Care Data, in defining pseudonymisation, itself states: *“In practice, pseudonymisation is typically combined with other anonymisation techniques.”*¹⁸

In the case of many releases of HES data, it has not been.

¹⁸ Pg 7 <http://www.isb.nhs.uk/documents/isb-1523/amd-20-2010/1523202010spec.pdf>

The “HSCIC Data Pseudonymisation Review – Interim Report”¹⁹ states: *“Pseudonymisation is not a “one size fits all” solution. There is a residual risk of jigsaw re-identification even when data items regarded as person identifiable have been replaced. Providing samples rather than whole dataset outputs and removing or obfuscating more information from the data lowers this risk.”*

Further, the NHS Anonymisation Standard states “essential requirements” for an output as being “Reduction in detail in indirect identifiers (such as date of birth, postcode)”²⁰ and has, as “essential standard requirements”:

- *“Post code truncated to either area code or district code*
- *No date of birth (e.g. transform to age, year of birth, or 5- year age band).*
- *No event dates (e.g. transform admission date to admission year, or month and year)”*

There have been and are releases of HES where none of those measures are applied, i.e. all data is still (re)identifiable. The data is only pseudonymised, which is described in the ICO’s Anonymisation Code of Practice as “a relatively high risk technique”²¹

Similarly, none of the measures listed in Annex 3 of the Anonymisation Code of Practice have been applied to the data. The data remains unprotected – no Statistical Disclosure Control has been applied at all.²² Even a single such release, as an unprecedented decision left unremarked, would have significant impacts on future interpretations of the Code and data releases.

While the Code of Practice is limited to safeguards in the Data Protection Act, the Government has committed HSCIC to go beyond this in satisfying fair processing for the nation’s medical histories, saying without caveat (emphasis added): *“The safeguards established by the Government – those in the 2012 Act and the announcement by my right hon. Friend the Secretary of State **that people could opt out of the collection and use of their data** – are welcome.”*²³

Risk of re-identification

It is understood that the protection measures expected of HSCIC are not supposed to be mathematically impossible to overcome. This is neither the required standard of the Data Protection Act nor the Anonymisation Code of Practice. However, an opt out that removed their data from onward dissemination would offer such certainty to those patients who wanted it. It is this crucial point that is been ignored by both HSCIC and the Department of Health.

¹⁹ 2.iv https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/401614/HSCIC_Data_Pseudonymisation_Review_-_Interim_Report_v1.pdf

²⁰ Section 6.4.2

²¹ p51 <https://ico.org.uk/media/1061/anonymisation-code.pdf>

²² <https://gss.civilservice.gov.uk/statistics/methodology-2/statistical-disclosure-control/>

²³ <http://www.theyworkforyou.com/whall/?id=2014-03-25a.49.0#g57.0>

The work of Professor Latanya Sweeney at the Harvard Data Lab has shown²⁴ how easily data can be reidentified. Taking US patient data equivalent to the Hospital Episode Statistics and information published in local newspapers, a project was able to reidentify 43% of individuals for whom the word “hospitalized” was used in a newspaper report. Information in newspapers – and on the web, and in social media – is “reasonably likely” to be in the public domain, yet continues to be entirely sensitive.

The particular identifier used to re-identify a record is beside the point. Identifiability comes from the fact that a single known event, such as a fractured elbow,²⁵ or a media notifiable condition,²⁶ especially where it is known to be the only operation of that kind in that hospital on that date,²⁷ provides the key to read that individual’s full hospital history. Indeed, transforming HES back into a complete, interrogatable event histories is what some commercial providers actively do.²⁸ We further note that whether linkage is facilitated via the NHS number itself, or an equivalent identifier or pseudonym derived from the NHS number, is entirely irrelevant.

The level of risk, which HSCIC itself refers to in some instances as “high risk”, is also arguably irrelevant. It is quite simply an unnecessary risk for anyone who has expressed dissent from having their data used in such a fashion. This would be most clearly illustrated in the case of a future data breach involving HES – of which there have been several in the past – where potentially millions of patients’ medical histories could be exposed to re-identification by third parties for a host of illicit purposes. Were the data of patients who had opted out to have been removed from HES, they simply would not be at risk.

While HSCIC may rightly claim its contracts forbid deliberate re-identification, contractual constraints can only go so far. Audit after the fact provides scant assurance; facts once discovered, even accidentally, cannot be unknown.

The ICO’s Anonymisation Code of Practice

At the launch of the Anonymisation Code of Practice, the Information Commissioner said that data that was not properly anonymised remained identifiable. We wholeheartedly agree with that assessment.

Despite its name, the Hospital Episode Statistics do not meet any of the normal definitions of “statistics”. HES data may be used to *produce* statistics, but they are not statistics in and of themselves. HES is raw data, not aggregated outputs.

²⁴ <http://dataprivacylab.org/projects/wa/index.html>

²⁵ <http://www.telegraph.co.uk/news/election-2010/7633714/General-Election-2010-Nick-Cleggs-wife-fractures-elbow-in-shopping-fall.html>

²⁶ <http://www.theguardian.com/world/2016/feb/28/ebola-nurse-pauline-cafferkey-discharged-from-hospital>

²⁷ <http://www.telegraph.co.uk/news/politics/ed-miliband/8666354/Ed-Miliband-undergoes-successful-nose-operation.html>

²⁸ <https://medconfidential.org/2014/commercial-re-use-licences-for-hes-disappearing-webpages/>

HES data is subject to no statistical disclosure control measures, including none of those means listed in Annex 3 of the Anonymisation Code of Practice. The ONS equivalent of such data would be the various survey microdata sets, which are subject to heavy statistical disclosure control prior to release. We are aware of no other broad microdata releases of administratively derived longitudinal data, let alone those absent data protections.

Which measures contained within the Code of Practice were and are taken, other than pseudonymisation and the issuing of a contract? In some instances, it appears none at all. Is this a valid interpretation of the Code of Practice for this data? Does the ICO believe that compliance with the Code of Practice requires only a contract?

“A wide range of clinical and administrative data about the hospital care episode are included in the HES record, such as event dates, and procedure/diagnosis codes, and these can also sometimes be used to reveal the identity of a person.²⁹ Different people may judge the sensitivity of health conditions differently, but hospital records can contain information that many people will feel is especially sensitive, such as a person’s HIV status, sexually-transmitted disease, mental health treatment, or a termination of pregnancy.”³⁰

“sending the data extract to be managed by the data recipient organisation according to terms set out in a data sharing contract. HES data may also be accessed indirectly through other data sources where HES has been combined with non-HES data to create a new data store”³¹

The latest Data Release Register from HSCIC shows that commercial re-users – who may pass data onwards to others – continue to be re-accredited to receive HES data.³² In fact, every information intermediary we referenced in 2015³³ has received re-approval in HSCIC’s Release Register.³⁴

The Anonymisation Code of Practice considers “consent”, as that is the general case. For data held by various parts of the NHS, the issue is for individuals who have actively taken up an offer from the Secretary of State to express dissent from their data being processed.

²⁹ On pages 2 and 3 of his report, ‘Review of data releases by the NHS Information Centre’, Sir Nick Partridge makes the point that clinical data in HES are codes and numbers and not “obviously personal descriptions of either patient or illness”, implying that as a result they are less revealing. See: http://www.hscic.gov.uk/media/14244/Sir-Nick-Partridges-summary-of-the-review/pdf/Sir_Nick_Partridge's_summary_of_the_review.pdf

³⁰ HES PIA section 4.4

³¹ HES PIA section 4.5 - disseminating from HES.

³² <http://www.hscic.gov.uk/dataregister>

³³ Note 5 - <https://medconfidential.org/2015/press-release-care-data-restart-announced/>

³⁴ HSCIC data release register, January - March 2016: <http://www.hscic.gov.uk/dataregister>

The legacy of care.data; ignoring alternatives

The original Privacy Impact Assessment for the care.data programme, published in 2014, showed that clinical data on the treatments of patients who had objected would continue to flow, and only their direct identifiers would be removed.³⁵ The Directions authorising this were deemed unacceptable by the Secretary of State, contributed to the pause in the programme that remains to this day, and led to the revocation of the original Directions and re-issuing of Directions on Patient Objections³⁶ explicitly stating that “no person-level or identifiable information about patients” should be collected where a Type 1 opt-out has been registered.

The dissemination of the HES data of patients who have objected is exactly the same issue. Patient have been told their data will not be shared, yet their data is being shared. This, despite the fact that HSCIC continues to develop “safe setting” solutions, which would mean that data did not have to leave HSCIC, and that HSCIC could provide strong guarantees to patients about exactly how their data has been handled, should appropriate standards be met on queries run against the data.

Data disseminated outside HSCIC is beyond the control of HSCIC, and is subject only to irregular audit and superficial examination. There is no way to know what queries were run, who ran them, or whether any subterfuge measures were taken.

As we said to the Health Select Committee in 2014, safe settings remain entirely appropriate, and in line with other administrative datasets made available for research. Contracts deeming millions of patients’ linked, dated, individual-level data “Anonymised in Context” are woeful by comparison.

What the public was told

Public statements are made not on the assumption of legalistic understandings of codes of practice, but much more broadly. Continued splitting of hairs between what is considered “anonymous” and what is “anonymised” contributes to ongoing public concern. Properly aggregated statistics are considered anonymous, since they do not involve individual-level data. HES comprises linked, dated, individual-level data and, despite its name, is not “statistics”.

The public context for the decision to rule HES “out of scope” of patient dissent includes the manifesto on which the current Government was elected, which stated (emphasis added):

³⁵ See penultimate paragraph on p9,

https://medconfidential.org/wp-content/uploads/2016/07/2014-01-15-care.data-PIA_1.0.pdf

³⁶

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/468439/patientobjectionsdirections.pdf

*“We will give you full access to your own electronic health records, **while retaining your right to opt-out of your records being shared electronically**”³⁷*

As a specific promise, and regardless of the opt-out mechanism (Type 1, Type 2 or other objection) this could not be more clear.

While it begins to fudge the issue, and introduces language open to interpretation, the NHS Constitution states:

“You have the right to request that your confidential information is not used beyond your own care and treatment and to have your objections considered, and where your wishes cannot be followed, to be told the reasons including the legal basis.”³⁸

NHS jargon may define “(personal) confidential information” in a particular way, but the most common public understanding of this is the clinical and other personal information held in a person’s medical records. Whether such medical information is or is not associated with the patient’s actual name or other details is rather less significant in many people’s eyes than the sensitive and confidential nature of the information itself.

Though it failed to reach many, and made no actual mention of a Type 2 objection, the junk mail leaflet sent in 2014 was unambiguous regarding hospital data (emphasis added):

*“Information from other places where you receive care, such as hospitals and community services, is collected nationally. **You should also let your GP practice know if you want to prevent the information from those places being shared.**”*

And numerous other Government statements in the ensuing period have been similarly unequivocal:

“the Secretary of State has already put in place an opt-out for patients who do not want to be involved in the process, which has not been the case in the past.”³⁹

“We have put in place the safeguard that people can opt out from having their data collected and used.”⁴⁰

“NHS patients need to know their data will be secure and not be sold or used inappropriately, which is why we have introduced tough new measures to ensure patient confidentiality.”⁴¹

³⁷ p38 <https://www.conservatives.com/manifesto>

³⁸ https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/480482/NHS_Constitution_WEB.pdf

³⁹ <http://www.theyworkforyou.com/whall/?id=2014-03-04a.240.0#g247.0>

⁴⁰ <http://www.theyworkforyou.com/whall/?id=2014-03-25a.49.0#g56.4>

⁴¹ <https://www.newscientist.com/article/2086454-revealed-google-ai-has-access-to-huge-haul-of-nhs-patient-data/>

“We will throw away these opportunities if the public do not believe they can trust us to look after their personal medical data securely.” ... “The NHS has not yet won the public’s trust in an area that is vital for the future of patient care.”⁴²

“To succeed, it is vital that the programme gives patients confidence in the way their data are used.”⁴³

This all being the case, we were astonished to learn, after the implementation of the Type 2 opt-out had been publicly announced, that patient objections are being deemed irrelevant for the dissemination of HES used for commercial purposes.⁴⁴

The continued issuing of commercial re-use licences means data continues to be passed to companies and organisations some patients have made it clear they absolutely do not wish to have access. And this practice is not mitigated in any way by attempts to claim uses are not “solely commercial” if an intermediary also services NHS customers.

So, despite all the promises, public statements and assurances since 2014, it is still the case that, for the dataset that cause the most public concern⁴⁵ – the very exemplar of what patients believed they were opting out of when they expressed a wish for their data not to leave HSCIC for purposes beyond their direct care – patients who believe they have opted out will still have their data included.

Summary

It may be that when HES was first created in the 1980s, pseudonymisation was considered sufficient to “anonymise” the data contained within it. This is clearly no longer the case, and was certainly not the case by the time the ICO’s Anonymisation Code of Practice was published in 2012.

The data dissemination policies and practices of HSCIC and its precursor bodies, when they came to light in 2014, were deemed wholly inadequate; the combination of pseudonymisation and contractual constraints they had operated did nothing to assuage public concern at the sale and sharing of patient data with commercial users and reusers.

Promises and public statements were repeatedly made that things would change. But when, in 2016, HSCIC finally began to act upon patient objections that had been ignored for two years, it chose to ignore all these promises and statements, and the most reasonable expectation of patients who had opted out, and instead continued with its previous practice: disseminating large volumes of linked, dated individual-level patient data to customers – in some cases, exactly the same customers that had caused public outcry in the first place.

⁴² <https://www.gov.uk/government/news/health-secretary-outlines-vision-for-use-of-technology-across-nhs>

⁴³ <http://www.theyworkforyou.com/debates/?id=2014-03-11a.279.4#g281.1>

⁴⁴ See enclosure number 2 - letter from HSCIC to medConfidential

⁴⁵ <http://www.hscic.gov.uk/article/7116/Consultation-on-the-Hospital-Episode-Statistics-Privacy-Impact-Assessment-Report>

Its justification? That pseudonymisation and contractual constraints were sufficient to meet all the required standards, i.e. despite a few updates to the contracts, nothing material had changed. From the patient's perspective, their opt out was simply irrelevant.

The dissemination of the data of those people who object to their data leaving the HSCIC for purposes beyond their direct care is *prima facie* wrong. It contradicts what those people were told. It is a betrayal of trust.

Hospital Episode Statistics data cannot and should not be considered "anonymous in context". In continuing to insist that it is, we believe HSCIC has failed to satisfy the Anonymisation Code of Practice and NHS Standards on Anonymisation (ISB 1523), not least because the richness of the information provided means it can be reidentified from information generally available, whether deliberately or accidentally by the recipient, or by any third party who gains access by lawful or unlawful means.

Wilful failure to address the nature and sensitivity of the actual data itself is unprecedented for datasets of this type. To allow HSCIC to continue to ignore standards and best practice, and even to hide behind the letter of the Code – especially in a case where a breach has been determined to have taken place – risks further erosion of public confidence not only in HSCIC and the handling of NHS patient data, but in the Code itself.

medConfidential