hscic Health & Social Care Information Centre

Draft Meeting Minutes

HSCIC Pseudonymisation Review Steering Group

Meeting 6, Wednesday 5 November 2014

Skipton House, London, SE1 6LH

Ratified by Steering Group January 2015

Attendees:	Role	Organisation
Kambiz Boomla	Observer	Confidentiality Advisory Group
Paul Croft (by dial in)	Senior Business Analyst	HSCIC
Harvey Goldstein	Academic expert on Data Linkage	University College London & University of Bristol
Xanthe Hannah (by dial in)	Observer	NHS England
Ian Herbert	Primary Health Care IT Specialist Group and GP Extraction Service Independent Advisory Group (GPES IAG) Member	British Computer Society
Julia Hippisley-Cox (by dial in)	Academic expert on Data Linkage and EMIS National User Group	Nottingham University
John Madsen (by dial in)	Head of Productivity & Efficiency	HSCIC
Hashim Reza	Consultant Psychiatrist	Oxleas NHS Foundation Trust
Jill Reeve (by dial in)	Project Support Officer	HSCIC
Chris Roebuck (Chair)	Director of Benefits and Utilisation and Review Co-ordinator	HSCIC
Matt Spencer	Pseudo Review Project Manager	HSCIC
Dave Wilby (by dial in)	Business Analyst	HSCIC
James Wood (by dial in)	Head of Infrastructure Security	HSCIC

Apologies

Paul Cundy	GP	General Practitioners Committee & BMA
Wally Gowing	Pseudonymisation Advisor	Observer

Alan Hassey	GP	IIGOP
David Ibbotson	Programme Head, care.data	HSCIC
Phil Koczan	GP	RCGP / Health Informatics Group
Geraint Lewis	Chief Data Officer	NHS England
Dawn Monaghan	Observer	Information Commissioners' Office
Nicholas Oughtibridge	Lead - Code of Practice for Confidentiality	HSCIC
Richard Pantlin	Social Care Representative	
John Parkinson	Observer	Clinical Practice Research Data Link
John Parry	Medical Director	TechUK
Daniel Ray	Head of Chief Information Officer Network	University Hospital Birmingham
Eve Roodhouse	Director care.data	HSCIC

1. Welcome and Introductions

1.1. Colleagues were welcomed to the sixth meeting of the HSCIC Pseudonymisation Review's steering group. The Chair thanked the group for its contributions over the past weeks.

2. Review of minutes/actions

2.1. The minutes of the 7th October Steering group were reviewed with comments raised as follows:

2.1.1. A member asked if the Interim Report, mentioned in section 2.3.4, had been changed. The Chair advised the Interim Report was published in July and that no further versions have been developed. It was further stated that the outputs from sub-groups were awaited and these would contribute to any recommendations made by the steering group.

2.1.2. A member, in referencing sections 3.3.3 to 3.3.7, stated that adding further terms to the S&T Vocabulary paper will mean it becomes too long. The Chair stated the purpose of the Vocabulary is for the Review's use only and its use had been agreed at earlier meetings.

2.1.3. It was stated there are variants of DSC in existence and as such there will be instances of DSCs authorising onward data sharing . The point made was DSCs are case specific.

2.1.4. It was agreed October minute item 3.4.5 be amended to state that Data Sharing Contracts specify whether any onward sharing of data is authorised and if so under what conditions.

2.1.5. A member asked what was meant by the wording 'after linkage' in 4.7 of October minutes. The Chair agreed that the paragraph should be reworded to "on identifiable and pseudonymised data".

2.1.6. The Chair stated there may be other examples of linking studies such as mental health that could be looked at.

2.1.7. A member referred to the CPRD audit which is looking at the possibility to pseudonymise data before extraction, rather than before linkage. The Chair responded that for any linkage studies they should not be looking to make it to specific and rule out other options.

2.1.8. The member is asking about the question, in 2.1.8, said the we are not just looking at pseudonymisation as it's not the only way of protecting data since other approaches such as sample selection could have a utility.

2.1.9. A member asked is it possible to pseudonymise data for specific purposes e.g. DOB only?

2.1.10.   Another member stated it needs to be very clear how data is treated whether using anonymisation of pseudonymisation.

2.1.11.   The Chair stated the Review has been set up to review pseudonymisation and as such should be looking at which fields should be pseudonymised and also where it is not appropriate. However other elements may be picked up when contrasting other approaches for impact when using pseudonymisation.

2.1.12.   The Steering Group agreed the review is looking at what data, that is received, processed and disseminated by the HSCIC, should be pseudonymised.

Action No. 1 – Project Manager to amend 3.4.5 of October minutes to reflect points raised in 2.1.4 and 2.1.5 and to re-word 'after linkage' in section 4.7 of October minutes as proposed in 2.1.6.

Action No. 2 – Project Manager, after making above amendments, to publish October minutes to the Review's website.

3. **Pseudo @ Source sub-group update**

3.1. The sub-groups chair, in referring to a progress report circulated to steering group, presented the groups Terms of Reference and provided a verbal update as follows:

3.1.1. The sub-group had met the previous week to finalise the ToR. There had been a number of revisions and recent inclusion of a work item to look at the impact of pseudonymisation at source on data sharing, data sharing contracts and agreements. The detail of these items would be worked on by Standards & Terminology sub-group but P@S sub-group will need to understand the detail when looking at pseudo at source as a model.

3.1.2. Andrew Dickinson, from HSCIC Information Governance, has joined the sub-group to work with the HSCIC business analysts Paul Croft and Dave Wilby. They are tasked with working up the detail for specific work plan deliverables on behalf of the sub-group chair. An update on progress on these deliverables will be provided at the sub-groups next meeting on 6th November.

3.1.3. The sub-group chair confirmed the approach to the group's deliverables by stating a final report will be sent to steering group in February 2015. The report will not be looking to propose a solution approach or technology but to define principles and requirements expected to be met.

3.1.4. Meeting being arranged with TechUK to hold an industry workshop event on12th December. List of prospective attendees being compiled, companies and organisations to be invited. There is a need to keep the numbers to around 40.

3.1.5. An agenda is currently in draft but likely to include following:

- What does the market offer?

- What does the industry feel about 'central', in the shape of HSCIC, being involved in pseudoing data that it receives,processes and disseminates...

3.1.6. There is a need to make sure we get a rounded view of Pseudo at Source in order to propose any recommendations to Steering Group and HSCIC EMT.

3.1.7. A member asked the sub-group chair whether any Patient groups had been identified for the workshop. For example 'Patient Partnership'.

3.1.8. It was noted that whilst there were no such groups currently planned to attend members were asked to provide details of specific groups to be represented to the sub-group chair.

3.1.9. The member, in asking about patient representatives' attendance, asked if payment of travel expenses is available. The Review's project manager will look into this and report back the member and group.

Action No. 3 - Project manager to confirm whether payment of expenses is allowable for patient representatives at P@S TechUK workshop event.

3.1.10.   It was noted that the Review is to look at public and patient engagement requirements in December once a number of other HSCIC patient stakeholder events are completed and as such no specific patient engagement consultation is currently planned by the Review.

3.1.11.   The member further asked if he could join the sub-group as a member. The sub-group chair responded by saying he would invite him to the next sub-group meeting.

3.2. Steering group members in responding to the sub-group chairs update and presenting of the groups ToR raised a numbers of questions.

3.2.1. The group's ToR mentions the three models of pseudonymisation, central, at source and a mixed model, there is in fact another that of the **Not to pseudonymise**.

3.2.2. A member in responding agreed with the comment but stated the review's work should be to clarify the position that identifiers in and what is to happen to these undertaken by referring to the Terms of Reference for Standards & Terminology sub-group has been sent to members for ratification at today's meeting.

3.2.3. In responding the member, in raising the comment in 3.2.1, said if this is correct then it needs to be detailed in the sub-groups ToR.

3.2.4. The sub-group chair responded by saying the group is looking at whether to use identifiable data. The group is look at the tools and processes to check the position use of identifiable data.

3.2.5. The member in responding to the chair's comment asked if the sub-group will be looking at all pseudonymisation approaches.

3.2.6. The chair replied that the group will not be doing a desk top analysis of all approaches but will seek to provide a view of different pseudonymisation approaches.

3.2.7. A member stated that Data Linkage & Data Quality sub-group is looking at deterministic approaches to pseudonymisation since the HSCIC does not currently extract data on names/addresses/free text and does not plan to do this.

3.2.8. The member, in asking what approaches are being looked at, stated that if there is not a detailed analysis, of pseudonymisation approaches, then the group would have a limited view of pseudonymisation.

3.2.9. The sub-group chair understood the concerns and points made, about pseudonymisation approaches, but the resources and limited timeframe available to the group meant it could not do the detailed analysis as expected. It may be that the group makes a recommendation, when it reports to steering group, to look at it further and to work with academia on the subject.

3.2.10. Another member agreed that the sub-group cannot undertake the detailed analysis of probabilistic methods (which are best suited for text data such as names addresses and free text), for when the review is supposed to report in February 2105, but it should be documented that this is something that will be required should there be projects which have a legal basis for extracting this type of data.

3.2.11. The member, who asked about what different approaches were being looked at, asked for clarification on what the other sub-groups were doing to ensure coverage in their work of the issue.

3.2.12. A member agreed clarification is required but thought only deterministic approach will be required however recognised that probabilistic approach needs to be looked at in future.

3.2.13.   The member, who raised the question on approaches, mentioned he has some notes on the different approaches that might be taken for linkages and these notes might be useful for the pseudo@source sub-group.

3.2.14.   Another member stated that if the sub-group, and Review's, work is limited to deterministic approach then it could say pseudo is not worth it without looking at other approaches.

3.2.15.   The steering group Chai in closing the discussion asked if the pseudo@source sub-group could look at the probabilistic approach in its ToR and to advise what resources and timeframes will be required to undertake this.

3.2.16.   The sub-group chair agreed to look at probabilistic approach and that it would be discussed at the group's next meeting. The chair stated it may have to scale other work items if the view was it should include it in its work programme.

Action No. 4 – Pseudo@Source sub-group chair to add detail to what extent it can support probabilistic approach in group'sToR and advise time line and resources to undertake.

3.2.17.   Other members, in reviewing the ToR, a number of comments were raised as follows:

- It was not clear what is being said penultimate para in section 2 of ToR.

- It was not clear what was meant by bullet point 7 in section 3.

- Members queried the reference to Mental Health and Prisons in final bullet of section 3. A member suggested the wording be changed to 'phased approach' to replace '*at the early stages ..*'.

- A member stated that pseudo algorithms are the same is each setting so this needs to be in the ToR.

- A member, in responding about Mental Health, made a general comment that MH data is no more sensitive than other care settings. Other members that there is a difference across care settings, as one episode in MH might be looked at differently to one in an acute setting. This would inevitably affect what data can be used.

- Another member commented that legislation differs for these different care settings

- The business analyst on the sub-group stated the need to look at a representative sample, of care settings, not just acute setting as this could give a skewed picture and the option may not be suitable for all settings.

- The steering group Chair commented that introducing too many things could make the process unmanageable. The sub-group needs to be clear on what it is looking and what can be delivered in the Review's timeline.

- The steering group Chair further stated that Mental Health linkage is key to fully understand any impact on Pseudo @ Source model.

- A member suggested that a workshop with EPR suppliers be considered.

- Another member stated that any workshop needs to be broader than EPR system suppliers.

Action No. 5 – Hashim Reza to provide a list of possible contributors to the workshop.

3.2.18.   It was noted the Pseudo @ source chair had left the call during the later stages of the discussion and that points made should be raised with him.

3.3 October's Action Log was reviewed with following comments:

- Action 15 – closed and September minutes approved and could be published.

- Action 23 – remains ongoing.

- Action 27 - remains open as sub-group work plans yet to be approved

4. **Data Linkage & Data Quality Sub-group update**

   4.1. The sub-groups Terms of Reference was presented to steering group members for ratification at today's meeting. It was noted that section 2, page 3 has been amended to resolve comments raised at an earlier meeting regarding inclusion of social care as a source for linkages without NHS number.

   4.2. The ToR was accepted by the steering group and should be published on the review's website.

   4.3. A member suggested adding John Sharp – HSCIC Head of Data Quality to the group as the work items covers data quality. This was agreed by the sub-group chair.

   4.4. An update on progress on the groups CPRD/HES study was given.

      4.4.1. The protocol for the study is under review with final comments expected by Friday 7th November. The sub-group chair said it involved complex datasets and it was important to ask the correct questions.

5. **Standards and Terminology sub-group update**

   5.1. The sub-group deputy chair provided an update on the group's deliverables.

      5.1.1. The Vocabulary paper, not formally presented at today's meeting is undergoing a number of revisions which will need to be presented to sub-group members for review. It is anticipated the paper will be presented to December's steering group meeting.

      5.1.2. The sub-groups ToR v1.0, circulated to steering group meeting, has undergone sub-group review in recent days with no comments received.

5.1.3. A member asked what the group was doing with respect to data Sharing Contracts as it was not clear in the ToR. The deputy chair committed to making the groups work on this topic more explicit in the ToR.

Action No 6 - Sub-group chair to provide a more explicit work description of Data Sharing Contracts (DSC) in the sub-groups ToR v1.0.

Action No. 7 - On completion of changes to circulate ToR to steering group, in advance of next mtg, for approval and publishing on the Review's website.

5.1.4. The sub-group chair presented the 'Context' paper which had been sent to steering group.

- The chair stated that the paper was proving difficult to compile with a number of revisions occurring as members had found the paper to be difficult to absorb.

- The paper is now about a third way though a further revision and as such it will require sub-group members review before presenting back to the steering group.

- The steering group Chair stated it should be a forward looking and not looking at what has happened in the past.

- The sub-group chair committed to work with the sub-group members and to deliver the paper to December's steering group.

6. **Work Plan review**

6.1. The steering group Chair presented the review work plan and asked for any comments.

6.1.1. A member mentioned that in December 2012 the care.data advisory group said it would look at Pseudonymisation being included in the programmes work. The member asked if this commitment is being kept.

6.1.2. Another member stated that HSCIC and NHS England (Tim Kelsey) told the Health Select Committee in February 2014, in response to a question from Barbara Keeley, that care.data would be looking at pseudonymisation in it's programme.

6.1.3. The steering group Chair respond that there is a need the Review to look at other programmes planning assumptions, including care.data, GPES and National audit programmes.

6.1.4. A member asked could the review look at extracting care.data in pseudonymised form. That is that it in extracted in Pseudo form and the Review looks into the results of this. This should be discussed with the member from a GP supplier.

6.1.5. A member commented that he thought it unlikely that GPES would undertake pseudonymisation at source for its extracts.

6.1.6. Another member said if care.data is using GP/HES for linking it will not be useful for the review as it will only be using deterministic matching approach. The member queried the suitability of this new extract for the Review's work. Another member disagreed with this comment since care.data have no plans to extract names and

addresses, and are currently planning deterministic linkage, then the review is highly relevant.

6.1.7. The Chair in responding stated there is a need to ensure the care.data Pathfinder programme is aware of what the Review is doing.

Action No. 8 – Review's project manager to ensure care.data Pathfinder programme is linked into the Review's work.

7. **AOB**

There were no items of any other business raised by members.

Next meeting: Wednesday 3rd December 13.00pm to 15.00pm at Skipton House, London.