Minutes of the meeting of the Confidentiality Advisory Group

07 May 2020 at Meeting via Teleconference

Present:

Name	Present	Notes
Dr Tony Calland MBE	Yes	CAG Chair
Dr Martin Andrew	Yes	CAG Member
Ms Sophie Brannan	Yes	CAG Member
Dr Patrick Coyle	Yes	CAG Vice-Chair
Dr Liliane Field	Yes	CAG Member
Mr Tony Kane	Yes	CAG Member
Professor Jennifer Kurinczuk	Yes	CAG Member
Mr Andrew Melville	Yes	CAG Member
Ms Clare Sanderson	Yes	CAG Alternative Vice-Chair
Mr Marc Taylor	Yes	CAG Member

Also in attendance:

Name	Position (or reason for attending)
Ms Katy Cassidy	Confidentiality Advisor
Mr Paul Mills	Senior Confidentiality Advisor
Ms Laura Gordon	CAG Assistant
Ms Natasha Dunkley	Head of Confidentiality Advice Service

1. Introduction, apologies and declarations of interest

2. Support decisions

Secretary of State for Health & Social Care Decisions

There were no applications that required a decision to be made by the Department of Health & Social Care senior civil servant on behalf of the Secretary of State for Health & Social for the **02 April 2020** meeting applications.

Health Research Authority (HRA) Decisions

The Health Research Authority agreed with the advice provided by the CAG in relation to the **02 April 2020** meeting applications.

THE HQIP decision has yet to be issued.

3. New applications – research

a. 20/CAG/0060 – The Health Improvement Network (THIN) database with primary care free text

Context

Purpose of application

This application from The Health Improvement Network (THIN) sets out the purpose of medical research which aims to enhance the current THIN database with free text data from general practice records, for use in research.

THIN is a long running research database which collects data from general practices records from across the UK for research purposes. The data is pseudonymised at source, with the GP practice having the key to be able to deidentify it. THIN currently stores coded data from approximately 12 million patients. This application seeks support for the creation of a free text database to enhance the current data it holds. In their submission the applicants propose using free text would enable researchers using THIN to be able to better understand doctor-patient interactions and decision making in primary care and provide greater depth of detail currently not recorded in the structured data.

Free text data would be extracted from participating practices to the THIN servers, where automatic deidentification would occur, removing 98% of any identifiers contained within the free text. The resulting free text would be stored until a request for use in research is made when, before release, two manual checks would be made on the data to ensure any remaining identifiers are removed.

A recommendation for class 1, 5 and 6 support was requested to cover access to the relevant unconsented activities as described in the application.

Confidential patient information requested

The following sets out a summary of the specified cohort, listed data sources and key identifiers. Where applicable, full datasets and data flows are provided in the application form and relevant supporting documentation as this letter represents only a summary of the full detail.

Cohort	All patients from GP practices providing data to THIN

Data sources	All GP practices providing data to THIN
Identifiers required for linkage purposes	No identifiers required for linkage purposes
Identifiers required for analysis purposes	Date of death Region of country Gender Ethnicity Occupation
Additional Information	Whilst no identifiers are required for linkage, the automatic tools used for de-identification of free text data are not 100% effective, so potentially identifiable information will be stored by THIN.

Confidentiality Advisory Group advice

The following sets out the Confidentiality Advisory Group advice which formed the basis of the decision by the Health Research Authority.

Public interest

The CAG noted that this activity fell within the definition of medical research and was therefore assured that the application described an appropriate medical purpose within the remit of the section 251 of the NHS Act 2006.

Members had a detailed discussion on the public interest of this study. Their unanimous view was that there was inadequate evidence of sufficient public interest

had been met. The applicants did not demonstrate the additional value that collecting free text data would provide to their existing data.

The free text data will contain highly sensitive information and will be stored on THIN systems in a potentially identifiable form for an indeterminate time, outside the control of GPs.

As such, the view of the committee is that any potential benefit of collecting the free text for use in research is vastly outweighed by the risk to public confidence if there were a release of such highly sensitive information in a potentially identifiable form.

Further, members noted that research findings indicate substantial public concern in research where there is a commercial organisation involved. Whilst this was not a primary concern of the committee, it was mindful of the impact of any inappropriate data release on public confidence in the use of patient information for research of this type.

Given THIN already collect coded data from practices, members requested reassurance from the applicant that no free text is being exported to THIN systems without a legal basis.

Data Controllership

Members consider that the data controllers in this application are the general practices. Given this, members were surprised that there was no attempt to provide evidence of engagement with general practices about the extraction of free text.

Members felt that, prior to release of any free text data, each partner of every practice should be fully involved in discussion, and sign up to, the proposal before the practice agrees to the release of such data.

Practicable alternatives

Members considered whether a practicable alternative to the disclosure of confidential patient information without consent existed in accordance with Section 251 (4) of the NHS Act 2006, taking into account the cost and technology available.

- **Feasibility of consent**

The committee agreed that consent would not be feasible.

- **Use of anonymised/pseudonymised data**

The committee noted the fact that the automated deidentification is 98% effective Given that there an estimated 12 million records within THIN, with each likely to have many free text entries, there is a real risk that an unknown percentage, potentially 100%, of patient record stored by THIN would have identifiable information within the free text records.

Further, members queried the accuracy of manual checks. Where a research request required many records (hundreds or thousands), members were concerned about the time taken per record to check and the accuracy of that check, and the resource available to be able to do this.

The committee also noted that the information will be stored for an indeterminate time by THIN, with no programme of deidentification of records.

The use of date of death in the context of this information was also a topic of concern by members, as patients can easily be identified by this information.

All these factors caused the committee real concern about the potential for release of identifiable, highly sensitive, information to researchers in both the UK and abroad.

‘Patient Notification’ and mechanism for managing dissent

It is part of the CAG responsibility to support public confidence and transparency in the appropriate sharing and use of confidential patient information. Access to patient information without consent is a privilege and it is a general principle of support for reasonable measures to be taken to inform the relevant population of the activity and to provide a right to objection and mechanism to respect that objection, where appropriate. This is known as 'patient notification'. This is separate to the local obligation to comply with the principles of the General Data Protection Regulation and Data Protection Act 2018.

The committee had strong concerns about the wrong and misleading information provided in the GP poster, where it states that the data will be anonymised, and all names and addresses will be removed. Nowhere does it state that the computer processing is not 100% effective, so does not give patients accurate information to make an informed decision.

Members also commented that the poster does not provide information to patients as how to find out more if they wish. The application mentions a leaflet, but this was not provided with the application.

In summary, the CAG felt patient notification processes were inadequate and misleading.

Patient and Public Involvement and Engagement

Meaningful engagement with patients, service users and the public is considered to be an important factor for the CAG in terms of contributing to public interest considerations as to whether the unconsented activity should go ahead.

Whilst the CAG noted that a lot of effort had been expended in undertaking Patient and Public Involvement and Engagement, it was also felt that much does not achieve the required aims. For example, there is little evidence that the fact the automated deidentification is 98% effective, as well as the number of records collected, so participants were unable to comment on this aspect.

Members also noted the concerns that participants clearly had, yet these do not seem to have been addressed in the application.

Confidentiality Advisory Group advice conclusion

In line with the considerations above, the CAG agreed that the application should be rejected.

Reasons for rejection

The below points summarise the reasons for rejection by the CAG:

The CAG does not believe the applicants have made a case that including free text data will produce outcomes that are in the public interest

This is especially evident when weighing the public interest and benefits of what is proposed versus the substantial risk of highly sensitive data being extracted from general practices.

General practices are data controllers in this application, yet no evidence was provided of any engagement with GPs. The CAG would expect, prior to any release of free text data, GP partner oversight and agreement of the activities proposed.

The 98% effectiveness of the automated deidentification process poses issues given the volume of records being collected. Due to the number of entries in the record of each patient, 98% effectiveness could mean that 100% of patient records are identifiable.

The patient information states "*The data in THIN are anonymised, which means that details that could identify a person such as names and addresses are not included. Words and sentences in health records are processed by a computer to remove any that might identify a person*". This is clearly wrong and inadequate, given the automated deidentification process is only 98% effective.

There is the potential for a large impact on public trust given the claims information is anonymised, but where there is a potential for information to be released in an identifiable form to researchers in both the UK and other countries.

Research shows that there is a large public concern of research when there is a commercial organisation involved. Whilst not the primary concern of this committee, the CAG notes the concerns the public would have if it supported the release of potentially identifiable data.

The PPI undertaken showed that participants did have concerns about the release of patient identifiable information in free text, and the applicant has seemingly not addressed these concerns in the application.

The use of date of death is problematic for the committee, given the ease of which that person could be identified with that information

The committee requested assurance that THIN are not already receiving free text data through its current processes.

The committee considered whether the current processes undertaken by THIN have a suitable legal basis under the Data Protection Act 2018, given current data is pseudonymised, not anonymised.

b. 20/CAG/0051 – Exploring the impact of mental health services in reducing suicide risk for those accused of online sexual offences against children within the Cleveland area

Context

Purpose of application

This application from Tees, Esk and Wear Valleys NHS Foundation Trust (TEWV) sets out the purpose of medical research which aims to compare the outcomes of people arrested on child sex offences between those who accepted or declined the involvement of mental health services, following their initial arrest.

Suicide prevention is seen as a national priority within the United Kingdom, with reduction in the numbers of suicides being part of Government policy and subject to regular report. Within the group accused of online child sexual offences, known as child sex offenders (CSO), the rate of suicide is higher. Estimates of increased risk of suicide in the CSO group range from 183 to 230 times higher when compared to the general population, with the greatest risk of suicide between 48 hours and one month of disclosure or discovery of the crime. Public benefit is highlighted as the National Police Chiefs' Council (NPCC) report regards increased suicide risk in this group, indicating the need to understand what measures may improve this outcome and understand whether the intervention carried out in TEWV is having a positive impact on suicide.

CSO cases matching the inclusion criteria will be identified using the Cleveland Police Database (approval obtained). The investigators will record name, date of birth and date of death (where applicable), as well as any other key demographic data available (e.g. yes/no answer to whether the individual is a professional) on an NHS laptop (or the police will send this data using an encrypted email). The researchers will link this with data on Tees, Esk and Wear Valleys NHS Foundation Trust PARIS IT system, in order to collect the data required for the research. Identifiable data will be stored on a separate spreadsheet to the study data, which will be identified purely by a study code. The resultant dataset will be used to highlight key factors that may make individuals arrested by the Cleveland Police more vulnerable to committing suicide and also highlight whether involvement of mental health services at an early stage is of benefit in this particular group. There is also the potential for service improvement, dependent on the outcomes of the research.

A recommendation for class 1 and 4 support was requested to cover access to the relevant unconsented activities as described in the application.

Confidential patient information requested

The following sets out a summary of the specified cohort, listed data sources and key identifiers. Where applicable, full datasets and data flows are provided in the

application form and relevant supporting documentation as this letter represents only a summary of the full detail.

Cohort	Any individual that was arrested by the Cleveland Police Paedophile and Online Investigation Team (POLIT) between 2015-2018. The sample size is estimated at 700-800 people.
Data sources	1. Tees, Esk and Wear Valleys NHS Foundation Trust
Identifiers required for linkage purposes	1. Name 2. Date of birth 3. Date of Death (where applicable)
Identifiers required for analysis purposes	1. date of arrest, 2. ethnicity 3. gender 4. aged between 40 and 60 years of age (yes/no) 5. married or residing with a partner (yes/no) 6. a parent (yes/no) 7. employed (yes/no) 8. not previously known to the police (yes/no) 9. suffering from a mental health issue (yes/no) 10. previously known to attempt suicide(yes/no) 11. participants live in the Cleveland area
Additional information	Due to the relatively small population size and the sensitivity of the work, the above may potentially be identifiable.

Confidentiality Advisory Group advice

The following sets out the Confidentiality Advisory Group advice which formed the basis of the decision by the Health Research Authority.

Public interest

The CAG noted that this activity fell within the definition of medical research and was therefore assured that the application described an appropriate medical purpose within the remit of the section 251 of the NHS Act 2006.

The committee agreed that this research has a clear medical purpose and would be in the public interest.

Scope

Members noted that information will be collected from Cleveland Police, which was outside the scope of s251 support since it did not relate to medical information.

Practicable alternatives

Members considered whether a practicable alternative to the disclosure of confidential patient information without consent existed in accordance with Section 251 (4) of the NHS Act 2006, taking into account the cost and technology available.

- **Minimising flows of identifiable information**

Members were somewhat unclear about the flows of data for this research, and commented that a data flow diagram would have helped the application. As such, members sought clarity that the cohort would first be identified through Cleveland Police and then linked with information from Tees, Esk and Wear Valleys NHS

Foundation Trust, not the other way around (in order to minimise the information accessed/collected from the Trust).

The committee also wished for clarity on the data being access/collected from the Trust. It was noted that the data collection was in the form of binary Y/N answers, but some of the research implies dates would be required/collected in order to answer these questions. The committee are mindful of the sensitivity of this data and potential identifiability of related dates and would like clarity on this.

Members also wished to question what identifiable data would be maintained for the duration of the study, as there was some uncertainty on this aspect.

- **Feasibility of consent**

All members agreed with the applicant's assertion that consent is not practicable for this population because some are deceased, others are in jail, and those acquitted may not wish to revisit this period of their lives.

- **Use of anonymised/pseudonymised data**

The committee noted that this study is not achievable with pseudonymised/anonymised data only given the need to link the two datasets.

'Patient Notification' and mechanism for managing dissent

It is part of the CAG responsibility to support public confidence and transparency in the appropriate sharing and use of confidential patient information. Access to patient information without consent is a privilege and it is a general principle of support for reasonable measures to be taken to inform the relevant population of the activity and to provide a right to objection and mechanism to respect that objection, where appropriate. This is known as 'patient notification'. This is separate to the local obligation to comply with the principles of the General Data Protection Regulation and Data Protection Act 2018.

The group were generally content with the notification, noting that it still needs formatting according to Trust policy. Members requested that a contact name was added to the poster, so that anyone who made contact can be sure that they are speaking to someone related to the research. The group also asked for the poster to be reviewed by the Lucy Faithfull Foundation for external comment (see more in Patient and Public Involvement and Engagement section below).

Members noted the intended locations for the poster to be displayed, but also commented that probation centres might be a useful addition to the locations.

Patient and Public Involvement and Engagement

Meaningful engagement with patients, service users and the public is considered to be an important factor for the CAG in terms of contributing to public interest considerations as to whether the unconsented activity should go ahead.

The group noted that Patient and Public Involvement and Engagement had not been undertaken for this application, and extensively discussed this at the meeting. Whilst Patient and Public Involvement and Engagement with the impacted cohort is usually expected it was agreed, exceptionally, that in this case it may not be useful. This is because part of the cohort is deceased/in jail, and any that may be available may not provide useful information.

However, the group noted the applicant had been in touch with The Lucy Faithfull Foundation. Members requested that the applicant asked the charity for their opinions on the research, as well as comments on the suitability of the poster.

Data Security

Members noted the sensitivity of the data being handled, as well as the consequences if this was inadvertently disclosed.

The group requested clarity on who would be accessing the information at Tees, Esk and Wear Valleys NHS Foundation Trust, and confirmation that all undertaking this activity will be doing so under a duty of confidentiality.

Members noted that the application is a trainee and is listed as the data custodian. As such, the group requested clarity on what would happen to the data if the applicant leaves the Trust. The group commented that either a change of data custodian or transfer of data out of the Trust would require an amendment

The group commented that the data will be collected on a laptop and requested assurances on the security of the device, including confirmation that it will be encrypted.

Confidentiality Advisory Group advice conclusion

The CAG agreed that there was a public interest in this activity, were supportive in principle of this activity proceeding, and therefore recommended to the Health Research Authority that the activity be provisionally supported. However, further information and actions would be required prior to confirming that the minimum criteria and established principles of support have been adequately addressed.

In order to complete the processing of this application, please respond back to all of the request for further information, and actions required to meet the specific conditions of support where indicated, within one month.

Request for further information

1. The applicant to speak with The Lucy Faithfull Foundation for comments on the proposed research.

2. Request comments on the poster by The Lucy Faithfull Foundation.

3. Include a contact name on the poster.

4. Provide a data flow diagram, detailing the flow of identifiable data in the research.

5. Confirm that identifiable information from Cleveland Police will be used to access information on the Tees, Esk and Wear Valleys NHS Foundation Trust IT system, not the other way around.

6. Provide a complete list of the data items being collected from Tees, Esk and Wear Valleys NHS Foundation Trust.

7. Confirm what data is being maintained by the applicant for the duration of the study.

8. Provide clarification on what will happen if the data custodian moves Trusts.

9. Confirm who will be accessing/collecting data at the Tees, Esk and Wear Valleys NHS Foundation Trust.

10. Confirm that all accessing data at Tees, Esk and Wear Valleys NHS Foundation Trust have a duty of confidentiality.

11. Confirm the security arrangements of the laptop to be used, including confirmation that it will be encrypted.

Once received, the information will be reviewed by a sub-committee of members in the first instance and a recommendation and decision issued as soon as possible. At this stage it may be necessary to request further information or refer to the next available CAG meeting. If the response is satisfactory and the outstanding actions listed in the specific conditions of support are met, a final support outcome will be issued.

Specific conditions of support (Provisional)

The following sets out the provisional specific conditions of support. These may change in the final outcome letter depending on the responses to queries.

1. Favourable opinion from a Research Ethics Committee. **Pending**

2. Confirmation provided from the IG Delivery Team at NHS Digital to the CAG that the relevant Data Security and Protection Toolkit (DSPT) submission(s) has achieved the 'Standards Met' threshold. See section below titled 'security assurance requirements' for further information. **The NHS Digital DSPT review for Tees, Esk and Wear Valleys NHS Foundation Trust was confirmed as 'Standards Met'. Email from NHS Digital received on 04 May 2020.**

c. 20/CAG/0056 – PEARL s251 Sensitive Health Records Template (Maternal smoking during pregnancy and intellectual disability)

Context

Purpose of application

This application from University of Bristol set out the purpose of medical research that seeks to investigate whether maternal smoking during pregnancy is associated with intellectual disability in children.

Tobacco use during pregnancy has been shown to influence foetal brain development and has also been associated with intellectual disability. It is unclear whether this link is causal or not. The applicants will use a variety of statistical techniques with the aim of improving understanding of linkage between tobacco use in pregnancy and intellectual disability in children. The smoking behaviour of mothers and fathers will be compared, and genetic methods used to determine causality. These analyses will be undertaken in several large population-based birth cohorts, including the Avon

Longitudinal Study of Parents and Children (ALSPAC) and sibling designs in other cohorts, to seek better understanding of the non-genetic causes of intellectual disability.

The applicants seek support to process confidential patient information from GP records, NHS Digital (HES data, the Mental Health Services Data Set) and data from the Avon and Wiltshire Mental Health Partnership, in line with ALSPAC's existing approval.

ALSPAC wish to:

- If the required sensitive data has already been collected through a previous project specific s251 support, repurpose the data for this purpose, in order to reduce flow of confidential information.
- For data not already collected though a previous project specific s251 support, request data from GP providers, NHS Digital and Avon and Wiltshire Mental Health Partnership using existing processes.

The data will then combined/anonymised by the ALSPAC team within UKSeRP (hosted by the University of Swansea but managed by ALSPAC staff). ALSPAC staff will then make the anonymised combined data available to the research team for this project.

A recommendation for class 1 and 4 support was requested to cover access to the relevant unconsented activities as described in the application.

Confidential patient information requested

The following sets out a summary of the specified cohort, listed data sources and key identifiers. Where applicable, full datasets and data flows are provided in the application form and relevant supporting documentation as this letter represents only a summary of the full detail.

Cohort	All people enrolled in ALSPAC, excluding those who have explicitly withdrawn from the study, declined consent to linkage to their health record, have not received ALSPAC

	fair processing information or have consented to data linkage.
Data sources	2. NHS Digital (HES and MHSDS data) 3. GP software providers 4. Avon and Wiltshire Mental Health Partnership NHS Trust
Identifiers required for linkage purposes	4. NHS Number 5. Date of birth 6. Date of death
Identifiers required for analysis purposes	12. Gender 13. Date of Birth will be used (by ALSPAC) to derive 'Age at Event' (expressed in days, minutes, seconds) and time intervals. This allows ALSPAC to provide information to researchers in event sequence order without disclosing Date of Birth or event date.
Additional information	Of the 15,000 ALSPAC participants, around 5500 have consented to data linkage and are not part of this request for support.

Confidentiality Advisory Group advice

The following sets out the Confidentiality Advisory Group advice which formed the basis of the decision by the Health Research Authority.

Public interest

The CAG noted that this activity fell within the definition of medical research and was therefore assured that the application described an appropriate medical purpose within the remit of the section 251 of the NHS Act 2006.

The group strongly agreed that the activity requested was within the public benefit, with the unique ALSPAC cohort able to provide insight into the research aims.

Scope

Members noted, as per the conditions of the original ALSPAC s251 support, approximately 5500 participants had consented to the activities, and so were outside the scope of this support.

The group commented that a number of participants in the cohort may have moved out of the Bristol area, and so would like to remind the applicants that s251 support is only for within England and Wales, and for the organisations for which support was requested (as above).

Practicable alternatives

Members considered whether a practicable alternative to the disclosure of confidential patient information without consent existed in accordance with Section 251 (4) of the NHS Act 2006, taking into account the cost and technology available.

- **Minimising flows of identifiable information**

Members noted the information requested and were content with the general processes used, given they have a precedent in previous applicants. However, the group were unclear on which identifiers were used under what circumstances and for which datasets. This was especially true for GP providers and Avon and Wiltshire Mental Health Partnership NHS Trust. Whilst a data flow was provided, members felt it was hard to determine the dataflows and which identifiers were used. As such, members requested an updated data flow diagram, indicating the flow of confidential patient information and pseudonymised data, as well as indicating what identifiers are used within each flow.

The application references community care data in a number of places. Members wished to clarify to what organisation these references mean, and whether data is being collected from any other sources.

The group wished the applicants to confirm that the collection of data does not include the collection of free text data.

- **Feasibility of consent**

The group were content that it was not feasible to obtain consent, at this stage, from remaining participants although noting that some effort was ongoing to consent participants were possible.

- **Use of anonymised/pseudonymised data**

Members agreed that the use of pseudonymised/anonymised data was impracticable, given the need to link datasets.

'Patient Notification' and mechanism for managing dissent

It is part of the CAG responsibility to support public confidence and transparency in the appropriate sharing and use of confidential patient information. Access to patient information without consent is a privilege and it is a general principle of support for reasonable measures to be taken to inform the relevant population of the activity and to provide a right to objection and mechanism to respect that objection, where appropriate. This is known as 'patient notification'. This is separate to the local obligation to comply with the principles of the General Data Protection Regulation and Data Protection Act 2018.

Members noted the patient notification materials and opt out information provided by the applicants and commended the applicants for displaying the notification for six weeks prior to data linkage. Members did notice that the fair processing materials indicated a project start date of 27 May 2019 and ask for the applicants to clarify/update this.

The group also noted that the notification states "*Remember: you need to let us know by the date given above if you do not want your records used in this way*." Members requested that this be changed to state that opt outs are preferred before this date, but participants can opt out at any time.

The committee thought that the participant notification should also link to the ALSPAC website opt out page so the participant can easily access the opt out information, and

should be clear that participants can opt out of this specific project only, not ALSPAC as a whole.

Members noted that the applicant stated that the National Data Opt Out would apply. Given this is not in place until September 2020, the group requested clarification whether this means the Type 1 opt outs will apply for GP records in the interim.

Patient and Public Involvement and Engagement

Meaningful engagement with patients, service users and the public is considered to be an important factor for the CAG in terms of contributing to public interest considerations as to whether the unconsented activity should go ahead.

Members were content with the Patient and Public Involvement and Engagement work undertaken by ALSPAC, noting investigations into linkage to sensitive data.

Confidentiality Advisory Group advice conclusion

The CAG agreed that there was a public interest in this activity, were supportive in principle of this activity proceeding, and therefore recommended to the Health Research Authority that the activity be provisionally supported. However, further information and actions would be required prior to confirming that the minimum criteria and established principles of support have been adequately addressed.

In order to complete the processing of this application, please respond back to all of the request for further information, and actions required to meet the specific conditions of support where indicated, within one month.

Request for further information

1. Provide a dataflow diagram which indicates the flow of confidential patient information and pseudonymised data, indicating which identifiers are used within which flow.

2. In the web notification, add a link to the ALSPAC opt out page.

3. Explain in the web notification that a participant can opt out of this specific project, without opting out of ALSPAC as a whole.

4. Clarify and update the project start date in the web notification.

5. Change the web notification to make clear that it is preferred for participants to opt out prior to the start date, but they can opt out at any time.

6. Define what the term community records mean, with respect to where data is collected from.

7. Clarify whether, in advance of the National Data Opt Out applying, whether the GP Type 1 opt out will be applied.

8. Confirm that the applicants are not collecting free text data.

Once received, the information will be reviewed by a sub-committee of members in the first instance and a recommendation and decision issued as soon as possible. At this stage it may be necessary to request further information or refer to the next available CAG meeting. If the response is satisfactory and the outstanding actions listed in the specific conditions of support are met, a final support outcome will be issued.

Specific conditions of support (Provisional)

The following sets out the provisional specific conditions of support. These may change in the final outcome letter depending on the responses to queries.

1. Favourable opinion from a Research Ethics Committee. **Confirmed February 2011**

2. Confirmation provided from the IG Delivery Team at NHS Digital to the CAG that the relevant Data Security and Protection Toolkit (DSPT) submission(s) has achieved the 'Standards Met' threshold. See section below titled 'security assurance requirements' for further information. **Pending. The NHS Digital DSPT**

review for University of Bristol, The Phoenix Partnership and the EMIS Group was confirmed as 'Standards Met' on the NHS DSPT Tracker (checked 10 May 2020).

NHS Digital DSPT review for Avon and Wiltshire Mental Health Partnership NHS Trust is outstanding. Please follow the security assurance requirements process below.

d. 20/CAG/0057 – PEARL s251 Sensitive Health Records Template (Incidence of psychosis and measures of psychotic experiences within the ALSPAC)

Context

Purpose of application

This application from University of Bristol set out the purpose of medical research to examine the proportion of children and adolescents with psychotic experiences (PEs) who transition to clinical disorder in adulthood, and estimate the extent to which they are identified by primary (PC)/secondary care (SC) services highlighting potential unmet need.

PEs are reported by 5-10% of the general population. Although usually transient, they are associated with increased risk of schizophrenia, but the natural progression of PEs and transition to schizophrenia in adulthood has not been examined in detail. Estimates of incidence are usually obtained from people presenting to clinical services and thus individuals who have not sought help would be missed from these estimates. Linking ALSPAC data with primary and secondary care records will allow the researchers to examine, prospectively, which early symptoms are most predictive of developing a psychotic disorder.

The applicants seek support to process confidential patient information from GP records, NHS Digital (HES data, the Mental Health Services Data Set) and data from the Avon and Wiltshire Mental Health Partnership, in line with ALSPAC's existing approval.

ALSPAC wish to:

- If the required sensitive data has already been collected through a previous project specific s251 support, repurpose the data for this purpose, in order to reduce flow of confidential information.

- For data not already collected though a previous project specific s251 support, request data from GP providers, NHS Digital and Avon and Wiltshire Mental Health Partnership using existing processes.

The data will then combined/anonymised by the ALSPAC team within UKSeRP (hosted by the University of Swansea but managed by ALSPAC staff). ALSPAC staff will then make the anonymised combined data available to the research team for this project.

A recommendation for class 1 and 4 support was requested to cover access to the relevant unconsented activities as described in the application.

Confidential patient information requested

The following sets out a summary of the specified cohort, listed data sources and key identifiers. Where applicable, full datasets and data flows are provided in the application form and relevant supporting documentation as this letter represents only a summary of the full detail.

Cohort	All people enrolled in ALSPAC, excluding those who have explicitly withdrawn from the study, declined consent to linkage to their health record, have not received ALSPAC fair processing information or have consented to data linkage.
Data sources	1. NHS Digital (HES and MHSDS data) 2. GP software providers 3. Avon and Wiltshire Mental Health Partnership NHS Trust
Identifiers required for linkage purposes	1. NHS Number 2. Date of birth 3. Date of death
Identifiers required for analysis purposes	1. Gender 2. Date of Birth will be used (by ALSPAC) to derive 'Age at Event' (expressed in days, minutes, seconds) and time intervals. This allows ALSPAC to provide information to researchers in event sequence order without disclosing Date of Birth or event date.
Additional information	Of the 15,000 ALSPAC participants, around 5500 have consented to data linkage and are not part of this request for support.

Confidentiality Advisory Group advice

The following sets out the Confidentiality Advisory Group advice which formed the basis of the decision by the Health Research Authority.

Public interest

The CAG noted that this activity fell within the definition of medical research and was therefore assured that the application described an appropriate medical purpose within the remit of the section 251 of the NHS Act 2006.

The group strongly agreed that the activity requested was within the public benefit, with the unique ALSPAC cohort able to provide insight into the research aims.

Scope

Members noted, as per the conditions of the original ALSPAC s251 support, approximately 5500 participants had consented to the activities, and so were outside the scope of this support.

The group commented that a number of participants in the cohort may have moved out of the Bristol area, and so would like to remind the applicants that s251 support is only for within England and Wales, and for the organisations for which support was requested (as above).

Practicable alternatives

Members considered whether a practicable alternative to the disclosure of confidential patient information without consent existed in accordance with Section 251 (4) of the NHS Act 2006, taking into account the cost and technology available.

- **Minimising flows of identifiable information**

Members noted the information requested and were content with the general processes used, given they have a precedent in previous applicants. However, the group were unclear on which identifiers were used under what circumstances and for which datasets. This was especially true for GP providers and Avon and Wiltshire Mental Health Partnership NHS Trust. Whilst a data flow was provided, members felt it was hard to determine the dataflows and which identifiers were used. As such, members requested an updated data flow diagram, indicating the flow of confidential

patient information and pseudonymised data, as well as indicating what identifiers are used within each flow.

The application references community care data in a number of places. Members wished to clarify to what organisation these references mean, and whether data is being collected from any other sources.

The group wished the applicants to confirm that the collection of data does not include the collection of free text data.

- **Feasibility of consent**

The group were content that it was not feasible to obtain consent, at this stage, from remaining participants although noting that some effort was ongoing to consent participants were possible.

- **Use of anonymised/pseudonymised data**

Members agreed that the use of pseudonymised/anonymised data was impracticable, given the need to link datasets.

'Patient Notification' and mechanism for managing dissent

It is part of the CAG responsibility to support public confidence and transparency in the appropriate sharing and use of confidential patient information. Access to patient information without consent is a privilege and it is a general principle of support for reasonable measures to be taken to inform the relevant population of the activity and to provide a right to objection and mechanism to respect that objection, where appropriate. This is known as 'patient notification'. This is separate to the local obligation to comply with the principles of the General Data Protection Regulation and Data Protection Act 2018.

Members noted the patient notification materials and opt out information provided by the applicants and commended the applicants for displaying the notification for six weeks prior to data linkage. Members did notice that the fair processing materials indicated a project start date of 07 December 2018 and ask for the applicants to clarify/update this.

The group thought the language in the notification was not in plain English and request the applicants to review and revise the notification accordingly.

The group also noted that the notification states "*Remember: you need to let us know by the date given above if you do not want your records used in this way*." Members requested that this be changed to state that opt outs are preferred before this date, but participants can opt out at any time.

The committee thought that the participant notification should also link to the ALSPAC website opt out page so the participant can easily access the opt out information and should be clear that participants can opt out of this specific project only, not ALSPAC as a whole.

Members noted that the applicant stated that the National Data Opt Out would apply. Given this is not in place until September 2020, the group requested clarification whether this means the Type 1 opt outs will apply for GP records in the interim.

Patient and Public Involvement and Engagement

Meaningful engagement with patients, service users and the public is considered to be an important factor for the CAG in terms of contributing to public interest considerations as to whether the unconsented activity should go ahead.

Members were content with the Patient and Public Involvement and Engagement work undertaken by ALSPAC, noting investigations into linkage to sensitive data.

Confidentiality Advisory Group advice conclusion

The CAG agreed that there was a public interest in this activity, were supportive in principle of this activity proceeding, and therefore recommended to the Health Research Authority that the activity be provisionally supported. However, further information and actions would be required prior to confirming that the minimum criteria and established principles of support have been adequately addressed.

In order to complete the processing of this application, please respond back to all of the request for further information, and actions required to meet the specific conditions of support where indicated, within one month.

Request for further information

1. Provide a dataflow diagram which indicates the flow of confidential patient information and pseudonymised data, indicating which identifiers are used within which flow.

2. In the web notification, add a link to the ALSPAC opt out page.

3. Review and revised the language used in the patient notification to be in plain English.

4. Explain in the web notification that a participant can opt out of this specific project, without opting out of ALSPAC as a whole.

5. Clarify and update the project start date in the web notification.

6. Change the web notification to make clear that it is preferred for participants to opt out prior to the start date, but they can opt out at any time.

7. Define what the term community records mean, with respect to where data is collected from.

8. Clarify whether, in advance of the National Data Opt Out applying, whether the GP Type 1 opt out will be applied.

9. Confirm that the applicants are not collecting free text data.

Once received, the information will be reviewed by a sub-committee of members in the first instance and a recommendation and decision issued as soon as possible. At this stage it may be necessary to request further information or refer to the next available CAG meeting. If the response is satisfactory and the outstanding actions listed in the specific conditions of support are met, a final support outcome will be issued.

Specific conditions of support (Provisional)

The following sets out the provisional specific conditions of support. These may change in the final outcome letter depending on the responses to queries.

1. Favourable opinion from a Research Ethics Committee. **Confirmed February 2011**

2. Confirmation provided from the IG Delivery Team at NHS Digital to the CAG that the relevant Data Security and Protection Toolkit (DSPT) submission(s) has achieved the 'Standards Met' threshold. See section below titled 'security assurance requirements' for further information. **Pending. The NHS Digital DSPT review for University of Bristol, The Phoenix Partnership and the EMIS Group was confirmed as 'Standards Met' on the NHS DSPT Tracker (checked 10 May 2020).**

NHS Digital DSPT review for Avon and Wiltshire Mental Health Partnership NHS Trust is outstanding. Please follow the security assurance requirements process below.

4. Chair's Report

The Chair's report was received and acknowledged by the group.

5. Any other business

No other business was raised.

The Chair thanked Members for their attendance and the meeting was closed.

Signed – Chair	Date

Signed – Confidentiality Advice Team	Date