Confidentiality Advisory Group

Minutes of the meeting of the Confidentiality Advisory Group held on 22 October 2026 via video conference.

Present:

Name	Capacity
Professor Lorna Fraser	CAG Chair (items 4a & 4b only)
Professor William Bernal	Alternate Vice Chair (items 4c & 4d only via correspondence)
Mr David Evans	CAG Member (Expert)
Mr Anthony Kane	CAG Member (Lay)
Professor Rachel Knowles	CAG Member (Expert)
Mrs Joanne Lloyd	CAG Member (Lay)
Mrs Sarah Palmer-Edwards	CAG Member (Expert)
Professor Sara Randall	CAG Member (Lay)
Cath Trigwell	CAG Member (Lay)

Also in attendance:

Name	Position (or reason for attending)
Mr Richard Reed	HRA Approvals Administrator
Dr Paul Mills	HRA Confidentiality Advice Service Manager
Ms Wangari Njiiri	HRA Approvals Administrator
Ms Caroline Watchurst	HRA Confidentiality Advisor
Professor Andrea Cipriani	Applicant for item 4a - Chief Investigator
Dr Karl Marlowe	Applicant for item 4a - Caldicott Guardian
Ms Helen Jones	Applicant for item 4a – Project Manager

Professor Geoff Hall	Applicant for item 4b - DATA-CAN Director
Johnny Chagger	Applicant for item 4b - DATA-CAN Data Protection and Information Governance Officer
Julia Millman	Applicant for item 4b - DATA-CAN Programme Director
Libby Cooper	Applicant for item 4b - DATA-CAN PPIE Chair
Nguyen Tran	Internal Observer - Cambridgeshire and Hertfordshire REC
Dr. Geraldine Boyle	Internal Observer - Social care REC

1. APOLOGIES FOR ABSENCE

Apologies for absence were received from: CAG Member Ms Kim Keltie. CAG Alternate Vice Chair Professor William Bernal sends apologies for the VC, but is chairing the modifications via correspondence.

2. DECLARATIONS OF INTEREST

2.1	**25/CAG/0112**	**IBD registry – modification to CAG**
	Conflict:	CAG Member Mr David Evans declared an interest in this item – as he is employed by NHS England, the data controller for the application. The Committee agreed that David should leave the meeting for the review of this application (ie, should not be included in the meeting via correspondence for this item).

2.2	**26/CAG/0109**	**DATA-CAN Cancer Database**
	Conflict:	CAG Member Mr David Evans declared an interest in this item – 4b, as he has previously worked closely with one of the applicants for the application. However that was some years ago and David has had no involvement with the DATACAN application. The Committee agreed that this did not constitute a conflict of interest, and David could continue to partake in the review of this application.

2.3	**26/CAG/0109**	**DATA-CAN Cancer Database**
	Conflict:	CAG Chair Professor Lorna Fraser declared an interest in this item – 4b, as she has previously worked on Connected Yorkshire with one of the applicants for the application.

		However that was some years ago and Lorna has had no involvement with the DATACAN application. The Committee agreed that this did not constitute a conflict of interest, and Lorna could continue to partake in the review of this application.

2.4	**21/CAG/0044**	**UK Longitudinal Linkage Collaboration (Study to NHS identifier flow v1)**
	Conflict:	CAG Alternate Vice Chair Professor William Bernal declared an interest in this item – 4d, as a relative is a senior researcher on the TEDS study, one of the LPS in LLC. However William has had no involvement with the TEDS application, or LLC application. The Committee agreed that this did not constitute a conflict of interest, and William could continue to partake in the review of this application.

3. SUPPORT DECISIONS

Secretary of State for Health & Social Care Decisions

The Department of Health & Social Care senior civil servant on behalf of the Secretary of State for Health & Social Care has agreed the advice provided by the CAG in relation to the **17 September 2026** meeting applications.

Health Research Authority (HRA) Decisions
The Health Research Authority agreed with the advice provided by the CAG in relation to the **17 September 2026** meeting applications.

Minutes:
The minutes of the following meetings have been ratified and published on the website:

- 11 September Precedent Set
- 17 September Full
- 1 October Full

4. NEW APPLICATIONS FOR CAG CONSIDERATION

4a	**26/CAG/0097**	**Count Me In - Informing patients about opportunities to participate in research**
	Chief Investigator:	Professor Andrea Cipriani
	Sponsor:	Oxford Health NHS Foundation Trust
	Application type:	Research
	Submission type:	New application

The Group reviewed the above application in line with the CAG considerations.

Applicants attended to discuss the application.

Prior to the meeting the applicants were informed that there were observers in attendance at the meeting. The applicants confirmed that they had no objection to the observers being present.

Summary of application

This application from Oxford Health NHS Foundation Trust (OHFT) is for the purpose of medical research enable OHFT to identify and contact patients in order to share information about research studies, prior to patient consent.

Traditionally, information about opportunities to participate in research have been shared with patients via their direct clinical care teams, however, using this approach, many patients have not been offered the opportunity to hear about research that may be relevant to them, due to clinical pressures. The aim of the Count Me In (CMI) initiative is to increase access to research by offering everyone the same opportunity to hear about relevant research opportunities, unless they have asked not to be included. This will be achieved using an 'opt out' approach to sharing information about research, allowing approved staff to identify patients who may be eligible for research studies and to share information directly with them. Patients are then able to decide for themselves whether they wish to receive this information.

This design will require 's251' support, as the individuals screening and approaching for consent are not considered to be part of the direct care team.

Confidential information requested

Cohort	Patients aged 18 and above, in receipt of care from Oxford Health NHS Foundation Trust, a community and Mental Health Trust, who have not opted out of receiving information about research opportunities. The cohort will be prospective from the point of implementation of the initiative, then both prospective and retrospective. Currently 141,666 unique patients (query run 6th September)
Data Sources	Oxford Health NHS Foundation Trust - Electronic patient records (Rio)
Identifiers required for purposes of screening for eligibility and contacting for consent	1. NHS number 2. Full date of birth (in order to calculate age) 3. contact details:

	4. patient's name (or a carer's name, in situations where a carer is listed as the first point of contact) 5. telephone number 6. email address 7. postal address
Identifiers required for analysis purposes	N/A – no specific analysis in this application, approaching for consent to research
Additional information	Deceased patients are excluded at the point the search is run (i.e. the code excludes any patient with a date of death present on the Trust EPR system when the code is run), so deceased patients should not appear in search results. Search results are refreshed every 3 weeks to ensure they stay up to date. This is confirmed again in the electronic patient record at the point of confirming contact details to ensure this is still accurate, prior to making contact.

Main issues considered, discussed and outcomes

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 section 251 of the NHS Act 2006.

Having reviewed the application and considered the risks and benefits involved, the CAG was also assured that the proposed activity was in the public interest.

The CAG requested confirmation on whether only adults would be recruited into the study. The applicant confirmed that this was the case and that, at that stage, the study would not include children or young people.

With regards to patient notification, the applicant was asked some queries prior to the meeting by the Confidentiality Advice Team (CAT). The applicant stated that the submitted poster is to signpost viewers to the website to find out more about the initiative. It does not contain any detail about the processing of data etc as this is outlined on the website. Applicants stated they are happy to develop an additional poster containing more processing details if recommended. The CAG recommended not developing an additional poster, but merely making any changes requested to the submitted poster.

The CAG identified that the National Data Opt-Out (NDOO) would be applied and noted that a project-specific opt-out had also been developed for potential patients. However, on the poster, there is no application specific opt out of the screening, and there is no statement to say the NDOO is respected. The leaflet and website also do not state that the NDOO respected. The CAG therefore requested that all patient notification materials be revised to include a statement

to say the NDOO is respected [Action 2a], and include further information on the poster on how individuals can utilise the project-specific opt-out process, including a phone number and email address [Action 2b]. The applicant confirmed that revisions to the patient notification materials were already underway and agreed to include additional information on the available opt-out processes.

The CAG clarified with the applicant that the opportunity to opt out would be before being approached to take part in any projects. The applicant confirmed this would be the case. The CAG also asked whether individuals who were approached and subsequently wished to object, would be able to opt out of future further contact. The applicant confirmed that, at the first point of contact, patients would be asked whether they were willing to continue the conversation regarding participation in the study, and patients would be able to opt out of further contact at that timepoint. CAG were content with this explanation.
The CAG commended the applicant for adopting a layered approach to patient notification. However, none of the materials currently explain that identifiers are processed by people who are not considered direct care team, without prior consent, with 's251' support from the HRA, following advice from CAG. The CAG therefore requested that the patient-facing materials make it explicitly clear that people who are not considered direct care team would have access to confidential patient information without consent, with 's251' support from the HRA, following advice from CAG [Action 2c]. In addition, the CAG requested that all revised patient notification materials be presented to a public involvement group to obtain feedback on the accessibility and clarity of the information provided [action 2d].

The CAG noted that the available evidence of public involvement was dated from 2022 and referred to commitments to ongoing public involvement activities. The CAG therefore queried whether any more recent public involvement had been undertaken. The applicant explained that continuous public involvement had taken place since then and that engagement with international collaborators had also been ongoing. The applicant explained the reason for documentation from 2022 being provided was due to them wanting to peer review and publish every step, and so the publications are a few years behind. The activity has been ongoing. CAG accepted this explanation.

The CAG noted that although the public involvement was extensive, there was some feedback indicating that some members of the public felt there should be a limit on the number of times an individual could be contacted if they did not respond to an invitation. The CAG therefore asked whether there would be a fixed number of approaches made to potential participants or whether this would vary depending on the study. The applicant explained that discussions within the public involvement groups had explored a range of circumstances and the limitations associated with inviting people to participate in studies. It was highlighted that some individuals might need to be re-contacted because they had been unable to respond to an initial invitation for reasons other than not wanting to participate, and that the team does not want to state/prescribe a maximum number of invites, as this might be considered gatekeeping. The applicant stated that the number of times potential participants could be

approached remained under discussion and that further advice would be welcomed. The CAG noted that, the number of invites for a specific study could potentially be 2 or 3 times, with no more approaches by letter, but that this might also vary depending on the individual study and its own REC approval. Whereas in addition, over a time period of 5 years, patients might be invited to participate in different studies on two or three occasions, for example. The CAG requested confirmation of the maximum number of times individuals might be approached to participate in projects, for each individual study, and overall [Action 3].

<u>Confidentiality Advisory Group advice:</u> Provisionally supported

The CAG was unable to recommend support to the Health Research Authority for the application based on the information and documentation received so far. The CAG requested the following information before confirming its final recommendation:

Number	Action required	Response from the applicant
1.	Support cannot be issued until a Favourable opinion from a Research Ethics Committee is in place.	
2.	The CAG request for the following changes to the patient notification materials and for updates to be provided to CAG: a) All patient notification materials be revised to include a statement to say the NDOO is respected b) include further information on the poster on how individuals can utilise the project-specific opt-out process, including a phone number and email address c) make it explicitly clear that people who are not considered direct care team would have access to confidential patient information without consent, with 's251' support from the HRA, following advice from CAG d) Ensure that revised patient-facing materials be presented to an appropriate public involvement group to obtain feedback on the accessibility and clarity of the	

	information provided.	
3.	Please explain if you have considered how many times an individual might be approached to participate in projects, for each individual study, and overall.	

The CAG also set out the following provisional specific conditions of support in addition to the standard conditions of support.

Number	Condition
1.	's251' support is provided for 5 years in the first instance, with a duration modification required to extend support.
2.	Confirmation provided from the DSPT Team at NHS England to the CAG that the relevant Data Security and Protection Toolkit (DSPT) submission(s) has achieved the 'Standards Met' threshold. **Confirmed:** The NHS England **25/26** DSPT review for **Oxford Health NHS Foundation** was confirmed as 'Standards Met' on the NHS England DSPT Tracker.

The Group delegated authority to confirm its final opinion on the application to the Chair and reviewers.

4b	**26/CAG/0109**	**DATA-CAN Cancer Database**
	Database applicant:	Professor Geoff Hall (organisation name not set)
	Database establishment:	Leeds Teaching Hospitals NHS Trust
	Application type:	Research Database
	Submission type:	New application

The Group reviewed the above application in line with the CAG considerations.

Applicants attended to discuss the application.

Prior to the meeting the applicants were informed that there were observers in attendance at the meeting. The applicants confirmed that they had no objection to the observers being present.

Summary of application

This application from Leeds Teaching Hospitals NHS Foundation Trust set out the purpose of creating a research database collecting data from partner organisations to support cancer research and improved patient outcomes.

The DATA-CAN Cancer Database will be a UK-wide research resource, designed to support the evaluation and improvement of cancer diagnosis, treatment and outcomes by bringing together routinely collected NHS cancer data from participating DATA-CAN centres into a single national data asset. The database will complement existing cancer datasets across the four UK nations by providing access to high-quality, clinically curated, multi-modal data that are not routinely available within national collections, enabling analyses at a scale that would not be possible through individual organisations or regional collaborations alone. Data will be securely transferred into the West Midlands Secure Data Environment (WMSDE), where the SDE technical team will undertake data quality assurance, linkage, de-duplication, pseudonymisation and other transformations necessary to create a research-ready dataset. The resulting database will be used to support approved research projects that demonstrate clear public benefit, including studies undertaken by clinical, academic and commercial researchers working in collaboration with the NHS. Project-specific pseudonymised datasets are then made available within secure Trusted Research Environments (TREs), to support research into cancer pathways, service delivery, treatment effectiveness, patient outcomes and the evaluation of existing and emerging therapies. The database is intended to generate real-world evidence to inform healthcare providers, NHS organisations, policymakers and regulatory bodies such as NICE and the MHRA, while also promoting collaboration, standardisation, transparency and reproducibility in cancer research through the preservation of data management and analytical records for future audit, validation and approved secondary research use.

Confidential patient information will be disclosed from participating trusts and health boards to the WMSDE, using either standard secure file transfer protocols or via a direct upload connection between the data provided and the SDE. The data will first be held in the Junction Zone of the Microsoft Azure Cloud, where it will be accessible to the data provider only. The data provider can manipulate data in this environment before it is transferred to the SDE Safe Haven. The Safe Haven will contain data from multiple data providers and is accessible only by the SDE technical team. The data will be pseudonymised before transfer to the Data Warehouse. Within the Data Warehouse, the data will be quality checked, made research ready and pseudonymised. The data is then transferred to the Staging Area, where study specific data is extracted before being sent to the TRE.

The applicant confirmed that all DATA-CAN data processing prior to this CAG application was not confidential patient information, and that the Flatiron application (23/CAG/0015) is completely separate and distinct from DATA-CAN.

Confidential information requested

Cohort	All patients with a diagnosis of cancer, who were aged 18 years and over at the time of diagnosis.
Data Sources	Routine NHS secondary care cancer datasets and associated clinical systems, alongside patient-

	reported outcomes collected as part of routine care, held by participating DATA-CAN hospitals.
Identifiers required for linkage purposes	• Name • NHS Number • Date of birth • Date of death • Postcode – unit level
Identifiers required for analysis purposes	• Postcode – district level • Gender • Ethnicity
Additional information	Items of confidential patient information to be held in data records within WMSDE. Full name NHS Number Date of birth Year of birth Postcode – unit level Gender Ethnicity

Main issues considered, discussed and outcomes

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 section 251 of the NHS Act 2006.

Having reviewed the application and considered the risks and benefits involved, the CAG was also assured that the proposed activity was in the public interest.

The CAG noted that DATA-CAN had apparently started in 2019, however no indication of any previous CAG application or REC application for this activity. The CAG also were aware of a 's251' supported project involving the same data controller and named applicant, which appeared to use a similar cohort of patient data - 'Flatiron Health UK Oncology Real-World Database', and were unclear how this new CAG application was connected to the already supported application, if at all. CAG asked if these projects were connected. The applicants responded that the only connection between the projects was that were both led by Leeds Teaching Hospitals NHS Trust, and both led by the database applicant for this study, but were otherwise fundamentality different projects that would never cross over. It was explained that Flatiron was a distinct project with data that was being released to a commercial company. DATA-CAN would also unite cancer data, but only within NHS governance, unlike Flatiron. DATA-CAN data would also facilitate access to the data, as a research database. DATA-CAN would be an entirely separate project to Flatiron, and that no DATA-CAN data can cross over for commercial usage to the Flatiron project.

In terms of the historical development of DATA-CAN from 2019, work had been conducted with a number of centres that brought data together, however the applicant assured CAG that this data was fully de-identified and did not involve access to confidential patient information without consent and outside the direct care team, and so previously no CAG application was required. The applicant also confirmed that the existing data in DATA-CAN would not be pulled into the SDE as part of this new application, as the data is already completed and archived. The CAG accepted these explanations.

The CAG asked for further clarification surrounding the description of the cohort. For example, how the cohort would be identified in the different centres that would submit into DATA-CAN. The applicants responded that not all hospitals with cancer patients would be asked to move all data on all cancer patients to the West Midlands SDE. Instead, specific data for a specific cohort would be requested per project, as per the general SDE model. Some hospitals had expressed a desire for the ability to send all of the relevant data to about all possible patients to DATA-CAN, as this was easier for them to manage. It was therefore important to offer a range of options for hospitals to provide data in a way that was best suited to them. The CAG asked for confirmation of the broadest conceivable cohort, including a defined time period for retrospective patients. The applicant confirmed the cohort was adult patient who were at least 18 at the time of diagnosis. The applicant stated that data from over the last 10 years was the most useful, but this might go back to historic data 15 or 20 years ago, as 5 and 10 year outcomes are of interest. The applicant is asked to confirm the maximum earliest date of diagnosis, for somebody to be included in DATA-CAN. This is important, so that people know whether or not they would be included in the cohort, if they wanted to opt out, and also because 's251' support is specific, providing a common law legal basis for accessing specifically defined data-sets. [Action 1].

The CAG requested confirmation that the patients identified as eligible from each Trust would only be patients treated for that particular episode of cancer at that Trust, and not just any patient with an historical cancer code within patient records. The applicant confirmed this, stating that although it would be important to hold the complete patient record to have confidence on past and current treatments, the submission of data form each site would not be derived from only a cancer code assigned to a patient record. The CAG accepted this clarification.

The CAG asked for further clarity on maximum number of patients included in DATA-CAN over time. The applicants responded that the average amount of data processed for a single project at an individual centre (Leeds) would typically amount to 500 patients. The applicant noted that there would be greater statistical power from larger cohorts, and that ultimately DATA-CAN would encompass tens of thousands of patients under broader criteria from multiple centres across the 4 UK nations. The applicant stated that the full database might contain 10-100,000 individual patients, however a specific project TRE may only contain 300-400 patients. The CAG noted that these numbers did not appear to align with the maximum cohort numbers in the

application, as it was stated that Leeds alone would be able to provide data regarding 349,000 individuals. The applicants clarified that actually Leeds has 480,000 potential individuals, but this would be only if all the data was sent on all patients. The intention would be to limit data transfer to specific projects, and therefore the number of patients would be determined by the number of projects using it.

The CAG asked if the identifiable data would remain in the WMSDE, or if the disclosures were single use. The applicants responded that the data would remain in WMSDE, and once in there, the aim is to be able to re-use it. Patients data would not be removed or deleted, but this would merely be added to over time. The cohort would therefore increase over time due to demand from projects. The CAG were not content with the applicants explanation regarding total cohort numbers, and asked for further clarity on maximum number of patients included in DATA-CAN over time, to which the applicant should provide the total maximum possible numbers of eligible patients in each site, rather than providing any estimates for specific projects. For example if Leeds have 480,000 potential patients, and the applicant is working with 20 sites total, this could be in the region of 9 million total potential patients. [Action 2]

The applicant is reminded that 's251' support only covers England and Wales. The CAG requested a list of every participating site in England and Wales. [Action 3].

The CAG noted that the project planned to retain identifiers in the landing zone of the WMSDE, and asked why this was required. The applicants confirmed this was required in order to link different patient episodes to the same patient, ie, it is important to be able to link all cancer recurrences together. The CAG accepted this explanation. However, CAG asked if it was possible to reduce these identifiers, for example in cases where a patient had died, as there would be no further updates to those records. The applicants responded that this had not been considered, but agreed that this could be factored in. Data may need to be held for up to 12 months in these circumstances due to any delays in receiving data after a patient death, but after that timepoint they would be happy to delete the identifiers of deceased patients. The CAG asked for this to be considered further and to confirm to CAG that confidential patient information could be removed from the dataset when a patient had died [Action 4].

The CAG noted that the project intended to collect full names and full addresses and asked if there was a reason for this amount of identifiable data collected. The applicants responded that NHS number, date of birth, sex and full postcodes would be needed to absolutely guarantee de-duplication. The applicant agreed that full address was not required for de-duplication, however did not confirm that it would therefore not be disclosed. The applicant did confirm that full address would not be needed beyond the landing zone, and this data would never be used in any analysis. The CAG asked for further justification for collecting full address. Additionally, the CAG requested written clarification that name, NHS number, date of birth, sex and full postcodes would be needed to absolutely guarantee de-duplication, and that this could not be done without name [Action 5].

The CAG noted that free text data would be retained in identifiable format, before being processed into the SDE, and asked how this was going to be processed? The applicants responded that the free text would be processed by WMSDE using a natural language processing (NLP) tool to remove any identifiers. Identifiers are required, to de-identify the free text, in order to establish what data would be required to be removed. The applicant stated that from experience this would be the most reliable way of removing identifiers from free text, and the plan would then be to make the full de-identified text available to any project that may require it. For instance, free text in pathology reports may be crucial data for some projects. The applicant stated that there are usually very low rates of error, however that there would therefore be a small percentage of patients free text data that could potentially contain confidential patient information. Due to the large datasets, low rates of error would still amount to a lot of patients. However, the applicant stated that this free text would not be freely available without specific justification from a project. The CAG asked if name and full address were only collected for the purpose of removal of identifiers from the free text data, and for a full justification as to why both of these data items were required. If name and full address are only used for this purpose, could these be deleted after the free text has been de-identified? [Action 6].

The CAG requested the actual expected error percentage of the NLP tool used to de-identify free text in WMSDE. The CAG also requested for information about examples of which types of identifiable data are slipping through. The CAG requested for further information regarding mitigations. For example, will there be any additional safeguards applied to the free text data released to applicants, to ensure the error rate in that particular data set is 0%? (for example, a check by WMSDE staff that the tool has de-identified the free text effectively in the dataset released to any researcher who requests free text). [Action 7].

The CAG noted that DATA-CAN has a core PPI group, for discussions on a weekly basis, however asked whether 30 public involvement contributors was considered adequate, especially as there was a strong gender imbalance with very few men, and lack of ethnic diversity within the group. The applicants responded that there was always an ambition to do more public involvement. The DATA-CAN core public involvement group provided input on the development of DATA-CAN. It was acknowledged that there was a gender and ethnic diversity imbalance, and this was hoped to be corrected. As well as the regular DATA-CAN core public involvement group, a larger meeting was planned, as currently public involvement regarding the national elements of DATA-CAN was lacking. The applicant also mentioned HDRS PPI and UseMyData, however this has not yet been undertaken. The applicant stated that would be happy to commit to much wider public involvement with a wider reach, and that the CAG application has changed the way they plan to undertake PPI, due to the response of the public with regards to questions asked relating to the requested 's251' support.

The CAG noted that DATA-CAN had some commercial funding and uses, but it

was not clear if public involvement consultations had explicitly discussed this aspect, and if any views had been raised as a result. The applicants acknowledged that cancer patients usually respect that commercial companies are involved when it comes to cancer treatment and care. It was also stated that commercial involvement had previously been discussed during previous iterations of DATA-CAN, however this can be carefully re-introduced as a topic for the PPI group.

The CAG asked for the applicant to confirm in detail exactly what public involvement work has been undertaken to date, including which groups of people, how they represent the cohort, and how many people were involved [Action 8a].

Further public involvement was also requested, with a larger group of people, better representing the relevant cohorts – i.e, a better gender balance, better diversity, and wider geographical representation. The use of confidential patient information without consent should be discussed, and any commercial input and uses should be discussed. Feedback should be provided to CAG [Action 8b].

Public involvement regarding the notification materials was requested [Action 8c]. The full set of notification materials should then be submitted for CAG review [Action 9]. Public involvement updates would also be required when submitting annual reviews (condition 1).

The CAG noted that one important principle of CAG support was the need to be explicit about whose data might be in the dataset, who has access to the data, and if people can opt out of their data being in this study. An individual letter and poster had been submitted and appeared to be in the early stages of devolvement in terms of transparency in this regard. CAG stated that there would be a need to work with public involvement to further develop the materials, and that a clear dissemination plan, for example through charities and other groups, as well as being displayed in clinical areas would be required, rather than planning to send individual letters out to everyone. The role of CAG should be described, as should the role of any commercial companies.

CAG asked how this would be managed to allow as many people as possible to access the patient notification materials. The applicants responded that the DATA-CAN public involvement group had been involved in the development of the materials, but it was acknowledged that more work was required to develop these. However, one of the hospitals involved has already completed a lot of public involvement work for this, and the team will liaise with them to see how this work could be applied at a local level. The CAG responded that final DATA-CAN documents would be needed, rather than differing information provided at each site. Consistent information is required across the documents that would be provided to all sites in England and Wales. This could be via a layered approach, with more detail provided on websites. The notifications should however be explicit about the use of confidential patient data without consent, any commercial purposes, what the database is used for, and clear options to opt out locally, prior to any data being disclosed outside if each Trust to DAT-

CAN. [Action 9a-i]. Further, a dissemination plan for these materials was requested [Action 10].

The CAG noted that the National Data Opt Out (NDOO) would be applied, but a project specific opt out would also be required. This should include a telephone or email address so that patients can contact to request project specific opt out, prior to their data being disclosed from the Trust to WMSDE. The applicants responded that there was a concern about how this would work. In discussions with a range of cancer centres, there were different views on how this should be done. Some hospitals already had preexisting processes, and did not want a DATA-CAN process to override local processes. Additionally, as letters would be sent about multiple projects, it could become very confusing to patients if DATA-CAN letters were also sent when patients had already been asked about their cancer data for a different project or purpose. As such, the proposal from the applicant was not to mandate a local opt out process if this was not desired by some hospitals. The CAG clarified that this would not involve an expectation to write to every individual, but to put information where it might be seen, such as a website, or at a suitable clinic. The CAG also stated that as the NDOO would block all secondary uses of data, it would be important to have local opt outs in place, so that a patient who wants to only opt out of DATA-CAN can do so, without also having to opt out of every other use of their data. The CAG asserted that some form of local DATA-CAN opt out would need to be in place. It may be that each centre will have a different mechanism to stop the use of data for the purposes of DATA-CAN. The applicant should ensure there is a clear way for a patient to request opt out via a contact telephone number or email address [Action 11]. The CAG additionally noted that if opt out mechanisms differ locally, it would be important to have standardised information so that everyone has information about what would happen to their data. This should state that the data would also go to West Midlands SDE, which was not apparent at present [Action 9b].

The CAG noted that the CAG form contained a statement that if there was a concern raised where harm may have been caused to a patient, this would be highlighted to the DATA-CAN team and to the local site, and that the patient may then be contacted, however the form also states that patients will not be contacted. The applicants clarified that if a patient was to be contacted, this would be at a local level and not by the DATA-CAN team contacting a patient. There would be a conversation with the chief medical officer at the local site to raise this concern, and would be passed back to the local site for them to follow up. The CAG responded that this proposal was fine, but the concern would be that if a patient was under the impression that no service would be able to contact them, it would be unclear how this issue had arisen, and as such should be present in the patient notification materials [Action 9i].

The CAG noted that it was clear that the project would involve the West Midlands SDE, but that there was also a possibility to extend to additional NHS sites. In this case, a modification would need to be submitted. Additionally, it was noted from the IRAS form that additional SDEs may be added to the project at a later time. Similarly, this should be submitted as a modification as and when this happens, and this would be especially important if the data was

to move to a different SDE (Condition 2).

The CAG noted that the proposed end date of the research database was 2031, but there was also references to an ongoing project with an indefinite end date. It was therefore asked if there was an exit strategy to 's251' support. The applicants responded that it was hoped that this database would become ongoing. However, it was recognised that renewals would be needed, and the current plan was to come back to REC and CAG in 5 years to request further renewal of the project. The CAG acknowledged this, however commented that a full renewal to CAG would not be necessary, merely a modification to extend the duration. (Condition 3).

Confidentiality Advisory Group advice: Provisionally supported

The CAG was unable to recommend support to the Health Research Authority for the application based on the information and documentation received so far. The CAG requested the following information before confirming its final recommendation:

Number	Action required	Response from the applicant
1.	Regarding the cohort, please confirm the maximum earliest date of diagnosis, for somebody to be included in DATA-CAN from retrospective historical data.	
2.	Please provide further clarity on the maximum number of patients included in DATA-CAN over time, to which the applicant should provide the total maximum possible numbers of eligible patients in each site, rather than providing any estimates for specific projects.	
3.	The CAG requested a list of every participating site in England and Wales.	
4.	The CAG request confirmation that confidential patient information would be deleted from WMSDE once a patient was deceased.	
5.	Please provide further justification for collecting full address. Please also provide written clarification that name, NHS number, date of birth, sex and full postcodes would be needed to absolutely guarantee de-duplication, and please justify	

	why this could not be done without name.	
6.	CAG asked if name and full address were only collected for the removal of identifiers from the free text data, and for a full justification as to why both of these data items were required for this purpose. If name and full address are only used for this purpose, could these be deleted by WMSDE after the free text has been de-identified?	
7.	The CAG requested the actual expected error percentage of the NLP tool used to de-identify free text in WMSDE. The CAG also requested for informaiton about examples of which types of identifiable data are slipping through. Will there be any additional safeguards applied to the free text data released to applicants, to ensure the error rate in that particular data set is 0%? (for example, a check by WMSDE staff that the tool has de-identified the free text effectively in the dataset released to any researcher who requests free text).	
8.	The CAG request the following in relation to public involvement: a) Confirm in detail exactly what public involvement work has been undertaken to date, including which groups of people, how they represent the cohort, and how many people were involved b) Further public involvement was also requested, with a bigger group of people, better representing the relevant cohorts – ie, a better gender balance, better diversity, and wider geographical representation. The use of confidential patient information without consent should be discussed, and any commercial input and uses should be discussed. Feedback	

	should be provided to CAG c) Further public involvement engagement is required for the notification materials. Please also see the following related action.	
9.	The CAG request that the full set of patient notification materials should be submitted for review once further public involvement has been undertaken using the updated versions, and any requested changes made. Please note that the updated notification materials should be explicit about the following: a) The use of confidential patient data without consent, ie, what data items are processed, and by which organisations. b) Explain that data will go to the West Midlands SDE. c) Include a statement to confirm that the HRA has given Section 251 support for the activity following advice from the Confidentiality Advisory Group d) Explain any commercial connections. e) Be clear about who the cohort is, and how far this will go back historically f) Explain what the purpose of the database is. g) List the options to opt out, with DATA-CAN specific opt outs having more prominence than the National Data Opt Out. A telephone number or email address should also be present for the patient to request DATA-CAN opt out, locally. h) The National Data Opt Out should be mentioned only as adhered to, rather than providing the link to how to opt out via this method. i) Note that, in rare cases, if a concern is raised where harm may have been caused to a patient, the clinical team j) may contact the patient directly.	
10.	The CAG request a dissemination plan	

	regarding the notification materials.	
11.	The CAG assert that some form of local DAT-CAN specific opt out is required to be in place, per local hospital Trust or centre, prior to data flowing to DATA-CAN. This will be suitable providing there is a clear way for a patient to request opt out via a contact telephone number or email address, and these should be present in the notification materials.	

The CAG also set out the following provisional specific conditions of support in addition to the standard conditions of support.

Number	Condition
1.	The CAG request ongoing public involvement ensuring diverse representation, and this feedback should be provided at annual review.
2.	The CAG note that the DATA-CAN project may expand to include additional NHS sites, as well as additional SDEs (or the movement of data to a different SDE). If and when each of these movements or expansions happen, a modification should be submitted to CAG.
3.	's251 support is provided for 5 years in the first instance, with a duration modification required to extend.
4.	's251' support is for English and Welsh sites only.
5.	Favourable opinion from a Research Ethics Committee. **Confirmed 13 August 2026.**
6.	Confirmation provided from the DSPT Team at NHS England to the CAG that the relevant Data Security and Protection Toolkit (DSPT) submission(s) has achieved the 'Standards Met' threshold. **Confirmed:** The NHS England **25/26** DSPT reviews for **University Hospitals Birmingham NHS Foundation Trust (host of West Midlands Sub-National SDE) and Leeds Teaching Hospitals NHS Trust (host of DATA-CAN)** were confirmed as 'Standards Met' on the NHS England DSPT Tracker.

The Group delegated authority to confirm its final opinion on the application to the Chair and reviewers.

4c.	25/CAG/0112	Inflammatory Bowel Disease Registry
	Contact:	Jenny Boatman
	Data controller:	NHS England
	Application type:	Non-research
	Submission type:	Modification

The Group reviewed the above application in line with the CAG considerations, via correspondence.

Summary of Modification

's251' support was originally provided to British Society of Gastroenterology (BSoG) for the purpose of establishing a national IBD Registry which would feed into national service development planning and fulfil national audit, IBD standards and quality improvement benchmarks; It was agreed that consent should then be obtained within a set time following the patient's entry onto the Registry, and it was suggested that this time period should be 18 months. A condition of support surrounding moving towards a consented model was applied; ***4. Support will be provided for a limited time period only whilst progress is made towards utilising a fully consented approach***.

Support to continue the non-research registry under 's251' support (rather than consent) has been extended in an ongoing fashion over the years, as the applicant has consistently reasoned that a consented registry is not feasible, however they have tried to meet CAGs requirements. In 2018, the application changed controllers from British Society of Gastroenterology (BSoG) to IBD Registry Limited. In 2020, the Chair supported an amendment to extend 's251' support, as Covid-19 had diverted efforts away from trying to establish the consented model. Other amendments have been supported since to extend, and in 2025 controllership once again moved from IBD Registry Limited, to NHS England.

The applicant currently only uses consented data for research, and therefore IBD Registry do not have an associated research application with CAG operating under 's251' support.

This modification aims to inform CAG of the change in lead applicant from Ceri Townley to Jenny Boatman.

This modification also aims to make changes to the original methodology. The final intention is that NHS England will move away from relying on 's251' support and use its powers already established under Directions, however the applicant has explained that this will take a number of years to implement, and the request for directions has not yet been submitted due to the upcoming changes which make the process uncertain. Although the data has been transferred from IBDR Ltd. to NHS England, further collection is on hold pending a review of the consent process as recommended by their Advisory Group for Data.

The main modification change is a **cessation of any consent model**:

CAG have received reports on the development of the consent model in annual reviews since the original support. NHS England is requesting that the requirement to move to a fully consented approach is dropped, and that the Registry processes confidential patient information entirely with 's251' support. The applicant has evidenced that this model is not working, as a low percentage of patients in the registry have consented, in the decade that it has been running. This is likely due to the capacity of clinicians, as consent would have been requested at Trust level. This is usually an accepted argument for registries to operate with 's251' support, rather than consent, as the requirement for consent would undermine the purpose of the Registry (as so few people would be included in the registry, such that benchmarking etc is not possible). NHS England propose to replace the current consent to join the Registry form with information about the s251 support, the National Data Opt-Out and Local Data Opt-Out, and materials have been provided.

With regards to research:

The applicant has never had or required a CAG research application, as they have undertaken research, or contacting participants about research only in the consented cohort. The applicant has stated that any dissemination of data from the NIBDR for research purposes will require the recipient to have an appropriate legal basis, or that the data are anonymous / pseudonymised (anonymous in context). This will be assessed by the established NHS England DARS process. However even anonymous/pseudonymous release would usually require a CAG research application, which the applicant does not currently have.

Applicants are also proposing to widen the scope of their non-research 's251' support to select and contact patients and gain their consent for the purposes of medical research, replacing the current consent to be contacted option on the registration and consent form. This isn't usually an option for non-research applications. The applicant has not detailed how they plan to contact the non-consented cohort, and CAG have not received a communications plan, or a proposed contact letter, or any public involvement about this element.

Main issues considered, discussed and outcomes

The modification requested was considered by the Confidentiality Advisory Group via correspondence, as part of the 22 October Full CAG meeting.

Overall, CAG were content to support the modification to remove the requirement for the IBD registry to seek consent from patients, and were content to advise support to cover the period until Directions are in place as an alternative common law legal basis. The public benefit of maintaining the registry for the purposes of national audit service development, quality improvement and patient safety justified.

However this will be conditional on updates to the notification materials, as the

Members found these quite limited in the information they currently provide. The notification materials need to make it clearer who the relevant cohort are, ie, which individuals might expect their data to be entered into the registry (e.g. Anyone who has been diagnosed by an NHS doctor? Do GPs enter patients onto the register? Who triggers the registry entry?). The material also needs to make clearer who is providing the data to NHS England.

The heading 'What has changed?' - implies this is intended for people already on the register or who know about the registry. For newly diagnosed patients, more relevant information to provide would be: you have been diagnosed, your data will go into the register for these purposes, you can find out more about the registry/NHSE management of your data here, you can object to this and withdraw your data by contacting x or y...). The CAG suggested that the applicants could consider an information leaflet/notification process that is suitable for both those on the register and those being added to it after the change in consent model.

The CAG agreed that the Public Involvement is supportive and satisfactory to support the change in consent model.

Regarding the proposed widening in scope to allow applicants to write to people to seek consent for research, the CAG noted that other registries that use their data for research purposes usually do not have any requirement to contact patients for their consent, although noted that this is clearly study specific, and the IBD may require contacting patients due to a requirement to physically participate in a clinical trial, or reply to a questionnaire study. Other non-research registries that use their data for research purposes, without consent, operate under a separate research 's251' support application. This includes anonymous/pseudonymous data extracts for research purposes, as the data has still been initially collected with 's251' support.

The applicant has not provided clarity on how this proposal, or any Public involvement regarding research use, or any detailed description of this in any notification, or any research application.

The CAG therefore advise that any research use of the data (anonymous/pseudonymised), or the contacting of patients to seek consent for research, or any other research use of the data collected with non research 's251' support, is not included in the non research support regarding this modification.

The CAG advise that if the applicant wants to use any data for research (even in anonymous/pseudonymous format), or approach patients for consent to take part in research, further extensive Public involvement should be undertaken, further changes to the notifications should be developed, including communications planning, and a new research database application for 's251' support should be submitted to CAG for review in a full CAG meeting. The applicant should approach the Confidentiality Advice Team (CAT) for advice regarding this submission if required. As part of this application, CAG do not generally require writing to all those currently on the registry who are not

consented for research, however if the methodology is appropriate and the public involvement supportive then direct contact for invitation to consider participation may be acceptable.

Confidentiality Advisory Group advice: Conditionally supported

In line with the considerations above, the CAG agreed that the minimum criteria under the Regulations appeared to have been met for this modification, and therefore advised recommending support to the Secretary of State for Health and Social Care, as below;

1. The modification, to;
    - change the lead applicant from Ceri Townley to Jenny Boatman,
    - and to remove the requirement for the IBD registry to seek consent from patients within 18 months after inclusion in the registry. The registry will operate using 's251' support to collect identifiers of all eligible patients, with 's251' support for the continued holding of the identifiers, to be used for non-research purposes,

   is supported, subject to compliance with the standard conditions of support.

2. The modification, to widen the scope of the non-research 's251' support to select and contact patients and gain their consent for the purposes of medical research, replacing the current consent to be contacted option on the registration and consent form, is **not supported**, on the basis described above.

Specific conditions of support

The CAG also set out the following specific conditions of support in addition to the standard conditions of support.

Number	Condition
1.	's251' support is in place for non-research purposes only. No 's251' support is in place for any research use of this data, including anonymous/pseudonymous release to researchers by NHS England. If the applicant wants to use any data for research (even in anonymous/pseudonymous format), or approach patients for consent to take part in research, further extensive Public involvement should be undertaken, further changes to the notifications should be developed, including a communications plan, and a new research database application for 's251' support should be submitted to CAG for review in a full CAG meeting. The applicant should approach the Confidentiality Advice Team (CAT) for advice regarding this submission if required.

2.	Please update the patient notification materials as follows and provide an update to CAG in 1 month; • Make the materials clear regarding who will have their data submitted to the IBD registry, and by whom. • Ensure materials address the likely concerns of both those whose data is already held, and those who will be newly included • Ensure adequate routes for opting out • Ensure adequate links to further information about the Registry
3.	Confirmation provided from the DSPT Team at NHS England to the CAG that the relevant Data Security and Protection Toolkit (DSPT) submission(s) has achieved the 'Standards Met' threshold. **Confirmed:** The NHS England **25/26** DSPT reviews for **NHS England**, **NHS Arden and Greater East Midlands CSU & Microsoft** were confirmed as 'Standards Met' on the NHS England DSPT Tracker **DHCW** has a valid Welsh IG toolkit.

4d.	**21/CAG/0044**	**UK Longitudinal Linkage Collaboration (Study to NHS identifier flow v1)**
	Chief Investigator:	Mr Andy Boyd
	Sponsor:	University of Bristol
	Application type:	Research Database
	Submission type:	Modification

The Group reviewed the above application in line with the CAG considerations, via correspondence.

Summary of Modification

This application has 'section 251 support' to create a pseudonymised research database - The UK Longitudinal Linkage Collaboration (UK LLC) - which will bring together data from UK longitudinal population studies (LPS), linked to data from national central datasets, into one secure 'Trusted Research Environment' (TRE), initially with a focus on COVID-19 research, and since expanded to wider purposes. Support is only required regarding participating studies which have current support under Regulation 5. Studies using a consented model do not require support. The LLC research database is UK wide but Regulation 5 support does not extend to Scotland or Northern Ireland.

The current 'section 251' support for LLC allows;

- the disclosure of confidential patient information from UK Longitudinal Population Studies (LPS) which currently operate under Regulation 5 support (named as NSHD, TWINSUK, ALSPAC, EPIC Norfolk & TEDS, & CaPS, with any additional studies requiring a modification to LLC application) to NHS Digital Health & Care Wales (DHCW) as a trusted third party.
- for DHCW to onwardly disclose identifiers to NHS England for the purposes of flagging patients as LLC participants and linking with relevant datasets, in order to supply the LLC database with pseudonymised linked data.
- and for DHCW to retain the identifiers relating to the LPS study cohorts to ensure de-duplication, and also to retain the key between the link ID and the key ID.

This modification sought 's251' support for the addition of the new NHS England National Disease Registration Service (NDRS) Cancer Consolidated Data Set as a data source.

This modification also sought 's251' to widen access to the UK LLC TRE to some international and commercial researchers. This has been supported by REC and is only regarding data that is effectively anonymous.

This modification also sought 's251' for the addition of new data flows/data sources/data processors regarding linkage to non-health administrative data. The specific support required is for UK LLC's trusted third party for linkage (DHCW) to send participants' identifiers (which are considered confidential patient information) to:

- the Office for National Statistics (ONS) for linkage to their Department for Work and Pensions (DWP) and HM Revenue and Customs (HMRC) records
- Department of Education (DfE) for linkage to their English education records.
- UK LLC is working with the other home nations to permit four nation linkage to education records.
- UK LLC's contracted data processors for place-based linkage (including the University of Leicester and City St George's, University of London) for linkage to place-based data. DHCW sends participants' address data (CPI) mixed in with real but randomly selected 'masking addresses' to They use the address data to model a number of environmental exposure estimates that are based on taking existing environmental sensor readings and, with other model inputs (e.g. traffic count data, weather pattern data), used to model pollution and other environmental exposures. UK LLC also collates and processes public domain datasets relating to a participant's address or neighbourhood (e.g. Lower Super Output Area) for inclusion in the UK LLC TRE.

The UK LLC Privacy Notice for LPS Participants explains how to opt out of each linkage or UK LLC altogether, ie. Patient notification has been updated. Public involvement has been undertaken.

The original outcome letter stated that no support was required for the non health sources detailed here, however this was because the applicant did not yet know how they were planning to link these in, and 's251' support does not cover non-health sources by themselves, as they are not confidential patient information. The CAG members reviewing at the time were aware and supportive of the future plans for non-health linkage. The applicant now has a plan in place for how to undertake the proposed linkage, and so have come to CAG with a modification to ensure the linkage with health data is covered. The processing of the non-health data itself is covered by the Digital Economy Act (DEA) (Data sharing across the public sector: the Digital Economy Act codes | ICO). UK LLC is accredited by the UK Statistics Authority as a processor under the DEA.

In addition, the only relevant applications at this present time (ie, an LPS that has 's251' support in its own right) that plans to allow linkage with any non health data are TEDS and ALSPAC (and ALSPAC already have 's251' support for these linkages). Linkages and research topics permitted for each LPS — UK LLC Dataset Documentation

These linkages are specific and the linked data will be deleted after each request as per policy.

Main issues considered, discussed and outcomes

The modification requested was considered by the Confidentiality Advisory Group via correspondence, as part of the 22 October Full CAG meeting.

It is noted that the applicant has stated in their query responses to the Confidentiality Advice Team (CAT) that: 'The flow of identifiers (where these are viewed as CPI) from NHS Digital Health and Care Wales (DHCW) for the non-health linkages is not undertaken under Section 251 support. Instead, the Digital Economy Act (DEA) provides the relevant legal gateway for these linkages and addresses the associated confidentiality requirements for approved research purposes.'

However CAG understand that not to be correct – the flow of identifiers from DHCW to those organisations is the only element that of this new proposed linkage that does require 's251' support. The DEA legislation provides a statutory gateway/power for the sharing and processing of non-health data by the newly added processors and the flow back to LLC. However 's251' support has been provided for the flow of CPI from DHCW to the listed non-health data processors, for linkage to the non-health data. 's251' support is not covering the processing of the actual non-health data.

CAG are overall supportive, and content to recommend support for this modification.

CAG noted that the LLC privacy notice provided (version 4.1, November 2025) does not include data access to international researchers for non-commercial purposes as proposed in this modification. The applicant should provide an updated privacy notice detailing both the commercial and international use.

LLC should also ensure updated patient information material/templates are provided detailing both the commercial and international use, for those organisations participating in the LLC under 's251' support, in order for them to inform their cohorts. These updates are usually provided as modifications by the participating LPS.

The applicant should also provide additional details of the nature of Commercial studies that will and will not be undertaken, in the form of some use cases covering what LLC see as commercial use, and examples of what commercial use LLC will not accept.

Confidentiality Advisory Group advice: Conditionally supported

In line with the considerations above, the CAG agreed that the minimum criteria under the Regulations appeared to have been met for this modification, and therefore advised recommending support to the Health Research Authority.

Specific conditions of support

The CAG also set out the following specific conditions of support in addition to the standard conditions of support.

Number	Condition
1.	Any research undertaken using linked non-health data, must have a medical purpose.
2.	Please provide an updated LLC privacy notice detailing both the commercial and international use, within one month.
3.	LLC should also ensure updated patient information material/templates are provided detailing both the commercial and international use, for those organisations participating in the LLC under 's251' support, in order for them to inform their cohorts. These updates are usually provided as modifications by the participating LPS.
4.	Please provide additional details of the nature of Commercial studies that will and will not be undertaken, in the form of some use cases covering what LLC see as commercial use, and examples of what commercial use LLC will not accept, within one month.
5.	Confirmation provided from the DSPT Team at NHS England to the CAG that the relevant Data Security and Protection Toolkit (DSPT) submission(s) has achieved the 'Standards Met' threshold. **Confirmed:** **The applicant must ensure that NHS England confirmation of 'standards met' for organisations processing confidential**

	patient information is in place. As there are more than 5 organisations, the DSPTs, or WIGTKs will not been individually checked by the Confidentiality Advice Team (CAT). This will be the responsibility of the applicant.
6.	Confirmation of a favourable opinion from a Research Ethics Committee. **Confirmed 16 July 2026**

5. ANY OTHER BUSINESS

There was no other business for discussion

Professor Lorna Fraser	02 November 2026
……	……..
Signed – Chair	*Date*
Professor William Bernal	02 November 2026
……	……..
Signed – CAG Alternate Vice-Chair	*Date*
Wangari Njiiri	23 October 2026
……	……..
Signed – Approvals Administrator	*Date*