Applications to use north east London Discovery Data:
Public Interest, Data Privacy and Access (PIPA) Review Policy
Approved by the North East London Discovery Board January 2020

1. Background

The Discovery Programme in north east London has been shaped with four main aims:

- To predict, anticipate or inform **individual** health needs from information or algorithms running in real time (or as near as possible) and to deliver the insight gained to the direct care team to enhance the patient's care record across the whole of their pathway, whether in primary or secondary care or elsewhere, thus creating the opportunity to improve or prevent adverse outcomes.

- To expand and improve the existing primary care informatics driven **population health** programme in east London, led by the Clinical Effectiveness Group at Queen Mary, to all health and care sectors.

- To enable the real time **reporting on programmes** by providers, commissioners and local authority public health supporting clinical improvement and new payment and organisational mechanisms. This would involve reporting on a pseudonymised or identifiable cut of the clinical data, as appropriate.

- To provide access to data for use by third parties such as commissioners, public health, academics and commercial/third sector organisations to **support research, service development and planning** for individual and public benefit whether on consented identifiable data, or the pseudonymised dataset. These uses form part of the benefits realisation anticipated for the Discovery Programme.

2. Objectives

The main objective of the PIPA review policy is to establish a process to ensure appropriate use of Discovery data in line with these aims, and with the expectations of and undertakings given to data publishers, patients and health care professionals contributing to the north east London Discovery Programme and the wider public interest (including being lawful and compatible with respect for human rights). The process is intended to enable informed decisions to be taken by the NEL Discovery Board whose members include data controllers.

The PIPA review process reflects the Discovery Programme's responsibilities and understanding of the legal basis for sharing Discovery data as well as the expectations of data controllers in agreeing to provide access to data, and of the public and patients. Access to and use of identifiable data not for the direct clinical benefit of a consented patient is a complex topic, reviewed in 2016 by

Dame Caldicott, National Data Guardian for Health and Care[1] and subject to the Data Protection Act 2018, the General Data Protection Regulation (GDPR) and other relevant legislation.

At all times, decisions to grant access will seek to ensure that any uses of Discovery data are consistent with its stated aims; protect patient privacy and confidentiality; have relevant scientific and ethics approval where appropriate, and make information publicly available about these uses, through the Discovery website and other routes.

It is intended that the PIPA Review Process is clear and transparent and implemented in a manner which is proportionate, accountable, fair and timely. The PIPA Review Process provides a framework for addressing and determining access; however, it is recognised that the nature of such requests may vary over the long-term in line with developments in the Discovery Data Service. Hence, they are designed to be flexible while complying with the principles set out above and will be subject to periodic review by the Discovery Board and the PIPA Review Panel and revised as appropriate.

3. Principles

Principles guiding access are aligned with recent national and international best practice and include:

- Use for all purposes in line with the aims of the Discovery Programme and that are in the public interest, without preferential or exclusive access for any person or subscriber. All subscribers, whether in NHS, universities, charities, government agencies or commercial companies, will be subject to the same application process and approval criteria.

- Applications to use Discovery data will be checked to ensure that proposals are consistent with the aims and expectations of data publishers contributing to east London Discovery programme, including, where appropriate, the relevant scientific and ethics approval.

- The anonymity and confidentiality of data will be safeguarded at all stages in the data life cycle. In particular, subscribers will enter into a legally binding Data Sharing Agreement with the north east London Discovery Programme and undertake not to make any attempt to identify patients, unless with the specific consent of individuals concerned, or with provision using some other legal basis such as section 251 support.

- Subscribers will be expected to access Discovery data hosted within an approved Trusted Research Environment and outputs of the data will be subject to independent review for statistical disclosure control purposes in line with best practice. Exceptions will be made where data linkage is carried out with explicit patient consent and considered exceptionally where a case is made for specialised facilities or functionality that cannot be provided within the approved Trusted Research Environment.

[1] Review of Data Security, Consent and Opt-Outs https://www.gov.uk/government/publications/review-of-data-security-consent-and-opt-outs

- Subscribers will meet the costs of the Discovery services for their approved purpose, including those costs incurred in providing access to the approved defined dataset, with a fixed charge for initiating the application review process and a variable charge depending on the complexity and scale of data required, whether any linkage is entailed, and the costs of hosting data in the Discovery nominated trusted research environment.

- It is expected that intellectual property arising from uses of Discovery data will be shared in ways that reflect the expectations and contribution of the publishers of data to the Discovery Programme as well as the contributions of the subscriber as set out in the data access agreement, and will not result in unreasonable restriction of access to resulting products or services by the NHS and related organisations.

- Subscribers granted access to the Discovery data will be required (as appropriate) to provide a summary of their approved project for publication in the public domain, to report on progress and impact, publish findings, attribute Discovery, and return data, code and derived variables as appropriate, so that they are available for others to use for purposes that are in the public interest. At the end of the period of access subscribers will undertake to archive their data within the approved Discovery data hosting facility with arrangements for certified destruction of data as appropriate in line with best practice.

- Discovery will seek to actively engage and communicate with patients, the NHS and our local communities on the uses of the Discovery data and the benefits that emerge.

4. Overview of Process

The application and review process for applications from subscribers is outlined in the **Figure** with an indication of the roles and responsibilities for each stage of the process.

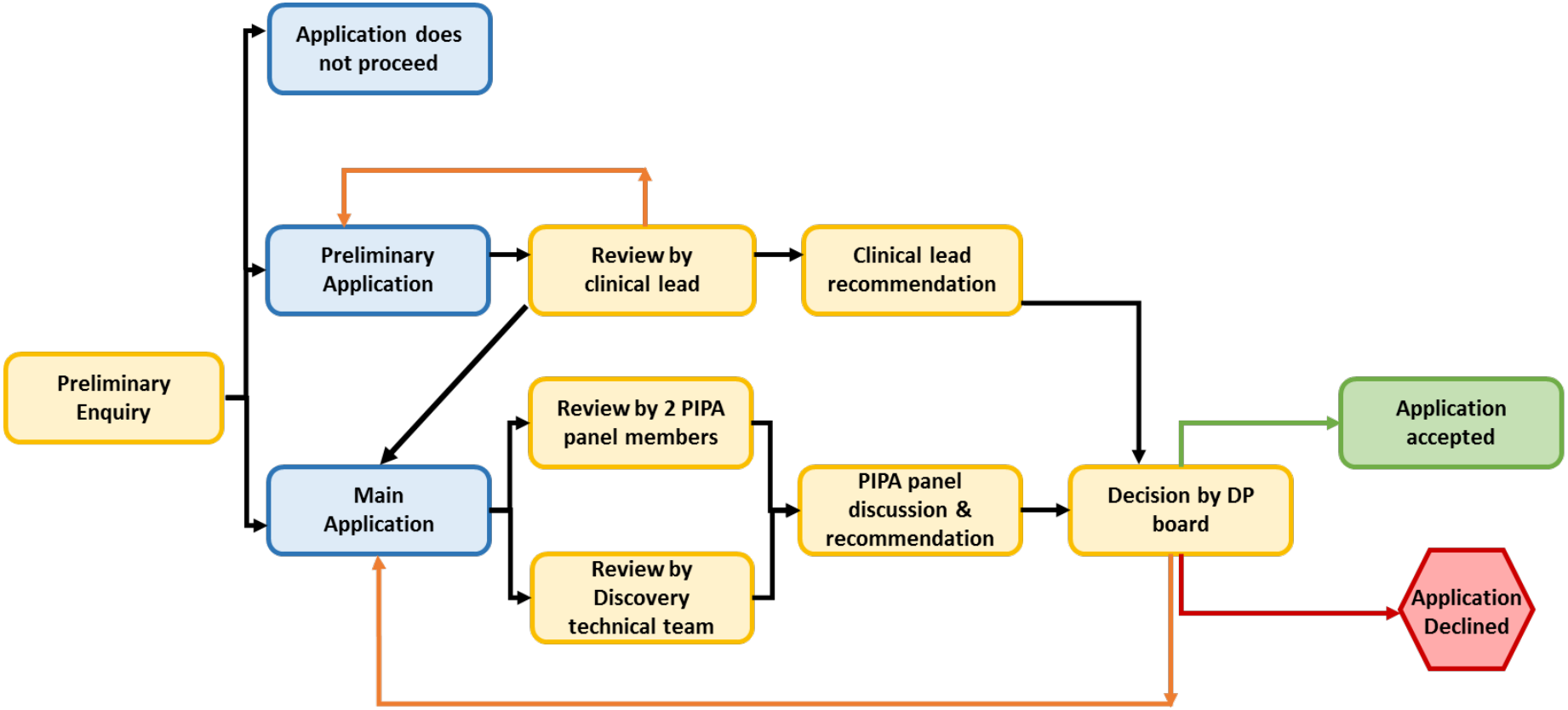

Figure Outline of application process for use of Discovery data

5. Application and Review process

Following a preliminary enquiry to the clinical leads to establish the feasibility of the proposal, subscribers will be invited to submit a preliminary application or to proceed directly to a main application (in the case of subscriber requests that are clearly articulated and/or there is a case for them to be fast-tracked).

a. Preliminary Application

The purpose of this stage is to

- Confirm the identity, status and organisation of each person applying or seeking to use Discovery data (the lead applicant and others who will have access to the data), and to confirm they would be appropriate and accredited to subscribe, and have access, to the data.

- Enable the applicant to understand whether the requested use is likely to be approved ie complies with principles set out above; whether Discovery contains appropriate and relevant data; and to provide an indicative cost of extracting and preparing data, and hosting data.

The applicant will be asked to supply:

- a short summary of purpose, objectives and scope for a non-specialist audience
- an explanation of how it meets Discovery Programme expectations and purposes
- a summary of requirements, to include patients, data, time period, geographical coverage, and additional data linkages
- an indication of the level and nature of NHS, public, commercial, industry or third sector involvement
- any anticipated IP considerations
- any ethical and information governance considerations, including whether any proposed use of identifiable data is with patient consent or has section 251 approval
- anticipated start date, duration of use and archiving or other longer-term considerations
- whether funding is available or to be sought, and likely source
- declarations of interest
- what fair processing and privacy statements are in place
- whether the applicant has conducted any public and patient involvement

b. Response to Preliminary Applications

Preliminary applications will be reviewed by Discovery clinical leads who may request further information regarding the purpose, feasibility, or other aspects of the proposal.

An initial response will be summarised, highlighting any potential issues, and once review is complete the applicant will receive this and one of the following responses: invitation to complete a main application with a note of any conditions or issues this will need to address; an indication that this is likely to be declined subject to final decision by the Discovery Programme Board

c. Main Application

Applicants invited by the Board to proceed to this stage will need to complete a main application form for consideration by the PIPA Review Panel and ultimately the Discovery Programme Board. This will include, as relevant:

- lead applicant and names of all those intending to access data (or if to be appointed, roles) including information on employing authority, IG accreditation[2]
- short summary of purposes: objectives and scope of proposed use for a non-specialist audience, updated as necessary
- rationale and background (including expected value/public benefit, aims and objectives, and, where appropriate, study design, analytic methods, power calculations, references);
- detailed data specification, including as appropriate cohort definition, codes/concepts, geography, time period, frequency of data extracts, and any additional data linkages
- an indication of the level and nature of NHS, public or third sector involvement
- any commercial or industry involvement and anticipated IP considerations
- proposed timetable (start; duration; key milestones and outputs)
- data hosting, archiving and destruction plans
- details of funding source, sponsorship, collaborators, and, if appropriate, any peer review (actual or proposed)
- details of any ethics approvals or assessments as appropriate
- letters of support or collaboration
- declarations of interest from applicants

[2] In line with specifications in data access agreement

d. Response to Main Application

Main applications will be reviewed by two members of the PIPA Review Panel as well as by the Discovery technical group: further information may be requested.

Their response will be summarised, highlighting any potential issues, and discussed at a meeting of the PIPA Review Panel who will make a recommendation to the Discovery Board.

The Discovery Board will discuss and issue one of the following responses:

- approval subject to completion of a Data Access Agreement and meeting data access costs
- approval as above and also conditional on specified issues being addressed within a specified time
- approval declined

Applicants whose proposals have been declined can opt to ask for reconsideration of the decision.

e. Reconsideration of Applications

If an application is declined either at preliminary or main application stage, applicants may request that their application be reconsidered by the Discovery Programme.

The process for having an application reconsidered is as follows:

- Within three months of the relevant decision, the applicant should submit a written request, giving reasons why they consider that the decision should be revised

- The PIPA Review Panel or the Discovery Board (as appropriate) will aim to consider it along with the original application (and any other information that it considers pertinent) and will respond to the applicant

- If considered necessary, the Discovery Board will seek additional advice from experts who are independent of the Discovery Programme and who have agreed to serve in this capacity; these experts will be asked to undertake a review based on documentary evidence but will not enter into direct discussion with the applicant

- If, following reconsideration, the application is declined then the applicant will not be able to submit the same proposal again within a 12-month period and then only if there is clear demonstration that the issues that led to approval being declined have been fully addressed.

6. Data Access costs

These comprise:

- A fixed charge for each application of £250 (plus VAT if applicable) to cover costs for the administration of the application payable on submission of the preliminary application form.
- A variable charge which depends on what is being accessed, calculated as a simple cost recovery of internal costs plus any third-party costs incurred by Discovery Data Services. This component will include charges for administration, refining data specification, data extraction and derivation and dataset preparation, data hosting and archiving, and post statistical disclosure control review. It will be payable prior to the release of any data or samples.
- Discovery will keep this charging policy under review (including the possibility of different levels of charge for different types of subscriber or content) to ensure that it continues to represent a fair, balanced and pragmatic approach and ensures that Discovery resources employed in providing data access are appropriately recompensed.

7. Data Access Agreements

- If the proposal is approved, then the Discovery Programme Manager will issue a Data Access Agreement to the applicant (along with a request for payment of data costs).

- Apart from inclusion of the specific details of the approved use (e.g. details of those having access to the data; the required data; the completion date for the project), the content of this Data Sharing Agreement, and the conditions contained within it, are non-negotiable.

- The Data Access Agreement will be considered to have been executed when the Discovery Programme Manager has received the Data Sharing Agreement signed by the lead applicant and their institution; and has received payment of agreed costs in cleared funds.

- Once the required data have been supplied to the lead applicant, the lay summary of the research project will be put on the Discovery Programme website (with any confidential information removed) in order that the public can see how Discovery data are being used. The code/concept specifications used by successful applicants will be made publicly available.

- In the event of any breach of the provisions of the Data Access Agreement by the subscribers, their employees or associates, licence to use Discovery data will be immediately revoked and may lead to other actions, such as informing the subscriber's institution and/or funders, as well as other regulatory bodies, and prohibiting further access from the applicant's institution.

8. Enabling Access to Data

- Data are to be hosted on data platforms that conform to the Discovery Programme's expectations in relation to the legal and ethical basis for data sharing and in line with current legal and ethical provisions.

- Projects using pseudonymised data will be hosted within an approved Discovery data environment, with access permissions created for approved and accredited users.

- Projects based on consented research studies can be hosted in other data environments provided these comply demonstrably with specified standards of cybersecurity and information governance as well as the stipulations of the Data Access Agreement.

9. Outputs from Discovery Data Uses

- For projects hosted in an approved Discovery data environment, clinical leads may undertake post analytic statistical disclosure control by reviewing outputs (tables, figures, graphs) from uses of unconsented patient level data before they can be exported from the data hosting environment. This policy will be kept under review in the light of practical experience.

- Subscribers are encouraged to report findings, with peer review as appropriate. While approval of such reports is not required from Discovery, the lead applicant must provide a copy of all of them to the Discovery Programme ideally in advance of first public presentation or publication in any format (e.g. meeting abstract, on-line report, paper journal). The applicant is also required to advise the Discovery Programme in advance if any report is reasonably likely to provoke controversy or reputational risks for data controllers.

- All publications and outputs should include the acknowledgement *"This project uses data about patients held by the north east London Discovery Programme which has been collected by NHS providers as part of their care and support."*

- Any reports or publications are required to be made open access.

- Within 6 months of publication or 12 months of when the approved application was to be completed, the applicant is required to archive data and any associated variables, codesets and coding as specified in the Data Access Agreement.

10. Intellectual Property Rights

- Discovery seeks to encourage use of the Discovery data for health-related purposes for public benefit.

- Discovery data publishers retain ownership of their rights in their data as supplied to the Discovery Data Service, while at the same time facilitating its development and use for clinical and other purposes and to support innovation and improvement in patient and population health and care.

- The Discovery Programme and its members retain all the intrinsic Intellectual Property Rights in the Discovery data (notably database rights and copyright).

- Subscribers are granted limited licences (but not any ownership rights) to use the data to conduct the approved project for a particular period of time. These rights are not assignable or transferable, and nor is there any ability to sub-license.

- If a subscriber creates innovations, additional data or variables as a result of their use, then Intellectual Property Rights in the subscriber-generated data will be shared by the subscriber, their institutions and the Discovery Programme in line with the Data Access Agreement and subject to the requirement to grant a non-exclusive licence back to the Discovery Programme for its use on an irrevocable, perpetual, worldwide, fully paid-up, royalty-free, fully sub-licensable basis.

- Data codesets, and new or cleaned variables, related coding together with appropriate annotation will be provided by the subscriber and archived in the Discovery GitHub or similar location and made available for use by others.

- The Discovery Programme will have no claim over inventions and associated IPRs that are developed by subscribers as a result of using the Discovery data, except as agreed in the Data Access Agreement and, notwithstanding the Data Access Agreement, unless such Intellectual Property Rights are used to restrict health-related uses and/or access to health-care unreasonably. In this context, the Discovery Programme does not seek to restrict applications for the protection of Intellectual Property Rights, rather their use in an unreasonably restrictive manner. In the event that conduct is considered unreasonably restrictive by the Discovery Board, it reserves the right to require that a licence of such rights is granted back to the Discovery Programme on an irrevocable, perpetual, global, royalty-free, fully sub-licensable basis for use by others or for public benefit more broadly. Discovery will keep this policy on Intellectual Property Rights under review to ensure that it continues to represent an equitable, balanced and pragmatic approach.