

DISCOVERY EAST LONDON

Interim Position Report – September 2017

1. BACKGROUND

In January 2016, a Project Initiation Document (PID) for the Discovery East London project was prepared, setting the objectives, deliverables, benefits and initial stage governance.

The Discovery Project is currently being implemented to establish, deliver and ultimately manage a secure data service with linked combined identifiable data from all systems supporting direct health care in east London. In its initial phase, the project has focussed on primary and acute care for the Waltham Forest, East London & City (WELC) CCGs and the Transforming Services Together (TST) Programme, but the aim is to move rapidly to include community health services, mental health and social care and to extend the geographic reach of the dataset to support the work of the full north east London sustainability and transformation plan – East London Health & Care Partnership (ELHCP).

Key

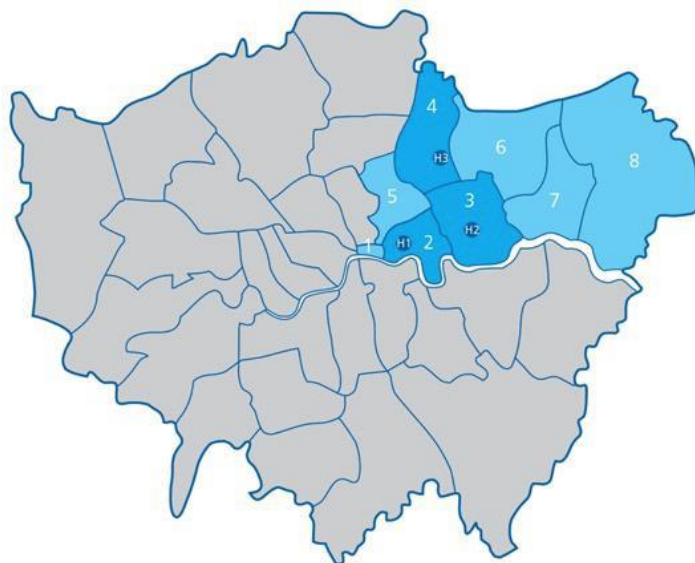
- Boroughs covered by *Transforming Services Together* (TST)
- Remaining boroughs covered by the Sustainability and Transformation Plan (STP)

Boroughs in the NE London Footprint

1. City of London
2. Tower Hamlets
3. Newham
4. Waltham Forest
5. Hackney
6. Redbridge
7. Barking & Dagenham
8. Havering

Hospitals with A&Es in the TST boroughs

- H1. The Royal London Hospital
- H2. Newham University Hospital
- H3. Whipps Cross University Hospital

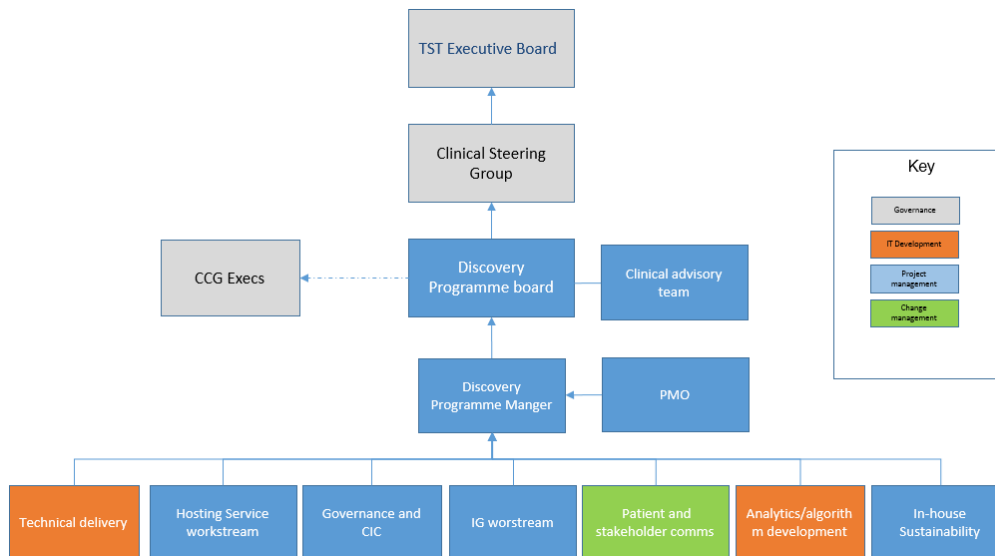


This project complements and enhances existing systems. It does not replace them and has four main aims:

1. To predict, anticipate or inform **individual** health needs from algorithms running in real time (or as near as possible) and to deliver the insight gained directly into the patient’s 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.
2. To expand 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.
3. To enable the real time **reporting on programmes** by providers and commissioners supporting clinical improvement and new payment mechanisms. This would involve reporting on a depersonalised or identifiable cut of the clinical data, as appropriate.

- To use data by third parties (commissioners, public health, and academics) **to support research, development and planning**, whether on consented identifiable data, or the depersonalised dataset. East London would thus become a research-enabled community.

The Discovery Programme governance architecture was originally shaped as below:



This Interim Position Report is structured in line with these original workstreams and seeks to provide a ‘year one’ status update against the objectives and deliverables of each of the seven workstreams identified, as well as more detail on the progress of the priority projects approved by the inaugural Discovery Board in November 2016.

2. TECHNICAL DELIVERY

In October 2016, Discovery achieved technical accreditation to access data from EMIS via the GP Systems of Choice (GPSoC) Assurance Steering Group (GASG) program. There have been a number of delays relating to contractual arrangements between EMIS and GASG. Funding is available to accredited subsidiaries until the end of the GPSoC-R contract in 2018 (and possibly beyond under a different mechanism). This will cover 2000 live GP practices for all accredited subsidiaries. NHS Discovery, Apollo and Informatica are accredited and 700 sites are live. We are currently working through the accreditation process to access GP data held on TPP (System One) and INPS Vision GP systems.

To access data prior to GASG accreditation, the program applied for GP data via the EMIS partner program. A contract was drawn up for EMIS to deliver daily extracts for one year. This service has not been reliable, resulting in non-payment of fees and a recommendation to novate the EMIS contract in favor of the GPSOC Service as soon as possible. It is expected that this will be complete by August 2017.

Primary Care

Though the extracts from EMIS have been unreliable, Discovery has had sufficient data from several practices to allow evaluation of the service, identification of bugs and issues and beginning data

checking tasks. As at 1st September 2017, the following numbers are available around practice sign up and switch on:

	Total Practices	Activated
City & Hackney	43	34
Newham	53	35
Tower Hamlets	35	35
Waltham For.	39	15
TOTAL	170	119

Discovery now has two subscriber-facing databases, containing identifiable data, being used for data validation.

Work with NHS Digital to support their SNOMED code mapping project has progressed. Discovery is receiving and transforming HL7 Admissions; Discharge and Transfer data from both Homerton Hospital and Barts Hospital Trust. The programme has informally expressed an interest in being considered as an early adopter for the SNOMED-CT work and early conversations to explore the implications of early adopter status are underway, although the project timelines are yet to be clarified by NHS England.

Development packages within the Discovery project continue with work on the Enterprise Query module linking into customer requirements and expectations, with feedback from CEG and the Hurley Group.

The programme has implemented a true single sign on solution across all products which, when combined with two factor authentication, delivers a secure efficient user experience.

Secondary Care

Discovery is now regularly receiving Admission, Discharge & Transfer (ADT) data from both the Homerton and from Barts Health. There are project plans in place with both secondary care trusts supporting delivery of both clinical data and clinical documents after the summer 2017.

Other Feeds

Discussions have begun with the East London Foundation Trust to explore **opportunities to share mental health data** with the Discovery platform. There has been a positive response from the Trust's Board and executive team and practical discussions to progress the development of infrastructure and technology links and associated use cases are scheduled over the summer 2017.

Discussions are underway with the **111/Out of Hours** services provided by Adastra to explore opportunities to feed this data into Discovery. These discussions are progressing in line with the current procurement process around 111/Out of Hours.

Agreement has also been reached with both City & Hackney and Newham to share **Local Authority National Child Measurement Programme** data with CEG and an initial set of weight data for 15,000 children across Hackney and approximately 18,000 children across Newham has been received. Initial review and analysis is underway. A similar agreement has been reached with Tower Hamlets and delivery of data from these feeds is being progressed.

Discussions have also commenced with Newham and the City of London to **shape use cases for linked health and social care data**, including broader council data such as housing. These discussions seek to optimise health and social care outcomes and broader public benefit from data sharing.

3. HOSTING SERVICE

Discovery East London is committed to protecting patient data, whilst making it available to meet the objectives of the programme. Over the past few years, there have been considerable advances in data hosting services and to maximise on this opportunity to deliver a more efficient and cost effective service, the programme needs to update the data hosting environment.

Data published into Discovery is controlled by strict data sharing processes and agreements. Those responsible for hosting the data are responsible for providing the technology to host and process the data store and do not access the data held within the store. Data is currently stored in a fixed hosted environment which is an N3 aggregated solution with IG level 3 accreditation to IGSO, ISO27001 standard and uses 2 factor authentications. Cloud services are made available to users on demand, via the Internet and N3, provider's servers as opposed to being provided by a company's own on-premises servers.

Access to the service by users, e.g. CEG and Hurley Group, was delayed by a requirement for local CSU teams to open ports. Access is now available and validated and, for future projects and relationships with CSUs, this access will be developed earlier in the roll-out plan.

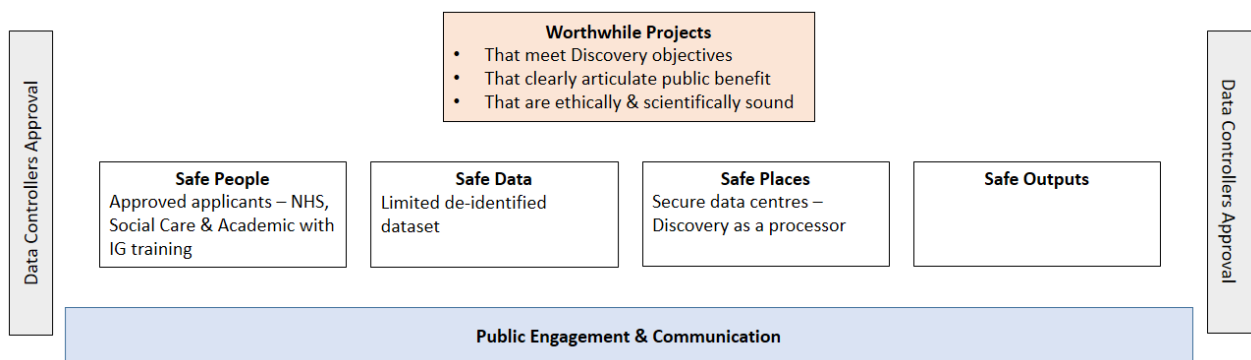
4. PROGRAMME GOVERNANCE

As part of its delivery, the Discovery East London project has established a formal mechanism to ensure that the service is managed by all contributing members to the benefit of all contributing members. In essence a joint Board, with clinical and management representatives (from all organisations that contribute data) governs both the project and the data governance. The holding organisation for Discovery is NHS Tower Hamlets CCG. The project office is within the CEG.

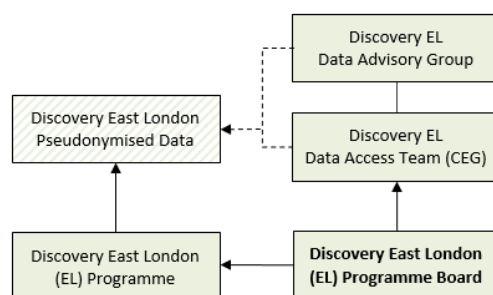
Later in 2017, work will commence to more build relationships and formalise the sustainable working model for the programme, initially focussing on bringing together representatives from the central contributors to Discovery to ensure that the purpose of the model is inclusive enough to remain fit for the future involvement of broader stakeholders (i.e. local authorities) whilst clearly articulating the specific benefits to be realised from that purpose.

5. INFORMATION GOVERNANCE

Access to and use of data by third parties 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 It is a requirement within the Discovery Programme to contribute to the establishment of a method of accrediting third parties for data access. In part, we have drawn on the experience of other large data repositories including those held by the Farr Institute, the Office of National statistics, UK Biobank and the UK Data Service. We have a firmly-established set of data sharing rules independent of supplier models to enable patient and population data sharing and have successfully used these to rules to control access to data for analysis. The governance and application review process is shaped in similar form to that used by the Farr Institute and National Services Scotland for data linkage for research purposes:



We have a tiered system with access for organisations contributing data, as well access by third party users such as University researchers and industry. We review applications to use the data against a set of criteria to achieve project objectives, to clearly articulate and measure public benefit and to ensure the ethical robustness of applications. The structure incorporates the following components:



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

Data Access Team (CEG)

Discovery East London operates as a platform within the QMUL Clinical Effectiveness Group (CEG) and therefore operates under its auspices to support improvement in direct clinical care across the East London system and research and development programmes agreed by CEG. CEG has an agreed role as *'trusted broker'* – not only will any agreed research comply with ethical and legal requirements, but it will not provide information for commercial advantage – *i.e.* commercial marketing of drugs, devices or other products. Discovery East London also operates within this framework and the CEG acts as the first phase of data access review when considering proposals to run algorithms or analysis on the data held within Discovery.

Data Advisory Group

The initial phase of data access review will then be supported by a broader local Advisory Group, advising the Discovery Board on uses of data from a local expert perspective. This Advisory Group's membership would include all relevant local expert stakeholders and be designed to effectively inform the Discovery Board on both matters of science as well as the specific views of the Discovery Data Controllers:

- GP Data Controller
- Hospital Data Controller
- Local Authority Data Controller
- At least one Patient & Public Involvement (PPI) Group member, supported by direct links with that PPI Group
- University
- External Academic

It is proposed that this Advisory Group would consider all uses of the data, including matters of relevance to Data Controllers such as performance and commercial uses. This group would have specific links with both the CEG's role as Data Access Group and with London-system international expert groups and specialist environments for data use (i.e. Trusted Research Environment for London).

Consent Model

Discovery East London is a data processor with exactly the same status as EMIS or RiO, for example. It only acts under instruction by the Data Controllers and would therefore be entitled to the data via instruction from the Trust. Within this model, the data would be shared with at least one end point of the Data controller's choice - the Discovery East London platform. Discovery East London is operating as a data processor under the control of the Data Controller and processing data on the Data Controller's behalf. The Patient Consent Model is relevant when determining the permission to share data with other Data Controllers utilising Discovery and would be implemented at that stage.

6. PATIENT & STAKEHOLDER COMMUNICATION

Significant work has been undertaken to ensure that initial stakeholders have been fully involved in early stage progress of the project. The project governance groups involve representatives from all direct stakeholders and contributors to the project, including those from GP practices, acute trusts

and mental health trusts. There are a range of information and engagement links in place to ensure that relevant interested groups are kept updated and informed of progress and to facilitate early stage development of specifications for data sharing and consideration of algorithms and analysis to be proposed for the data once in place.

There is considerable work to be undertaken to appropriately shape engagement and involvement with the public and patients directly. At this stage, data is shared with Discovery through data controller (i.e. GP and acute trust) consent and so direct engagement and communication with patients and the wider public has been confined to supporting GP practice staff with a set of 'Frequently Asked Questions' (see **Appendix A**) and to uploading an early Fair Processing Statement onto the Data Controller Console, as below:

The Discovery Project - This is a new project looking to create a new data service where the local providers of your care (like the GP practices and hospitals) will link up their data more seamlessly in order to provide you with better quality care. Your local health and social care providers all have their own IT systems which hold your information in a way that is not available to the other organisations providing your care. For example, GP data is not available to the hospital and vice versa. This service aims to pull all that data together to create a better overall picture of your care needs. One of the main aims of this project is to put you in charge of giving your consent for the use of all of this data so you can say who has access to what and where they can see it. This project is new and is not up and running yet but we will provide you with more information as it progresses and about how you can engage with us on it.

It is acknowledged that this early Fair Processing Statement is limited in terms of detail and will require more in-depth information to support ongoing patient and public engagement activities.

Current national work to operationalise the content of Dame Fiona Caldicott's third report provides structures for the use of depersonalised data and the engagement and opt out requirements to be undertaken in the use of this data. As Discovery East London's data is depersonalised at source, it is suggested there will be no national requirement to offer an opt out to local patients. However, the project is currently considering whether there is indeed value in testing that position with patients and the local population during engagement processes. This will also be considered in line with the recommendations of the Information Commissioner's Office (ICO) General Data Protection Regulation (GDPR) 2018² guidance, coming into effect on 25th May 2018.

Initial engagement with patients and the public is scheduled after the summer 2017, involving presentation at the Barts Health Open Day on 16th September 2017 and using the same materials with groups of patients involved in other projects across the geography. Local conversations with Patient & Public Engagement leads across the four CCGs are underway, including initial discussion with HealthWatch teams to explore the most productive routes to link with communities and populations which are not normally engaged with during early stage work. Discussions are also underway with various national leaders in involvement and engagement activities with patients and the public on data use (Citizens Juries CIC, the Farr Institute, Greater Manchester AHSN [DataWell],

² <https://ico.org.uk/for-organisations/data-protection-reform/overview-of-the-gdpr/>

the Wellcome Trust's Understanding Patient Data (UPD) programme with a view to building a cross-thematic agenda to test questions of data use, data users and the language required to most effectively involve and engage patients and the public in the improvement of their direct care and population health, comparable across national populations.

7. ANALYTICS / ALGORITHMS IN DEVELOPMENT

Aligned with the fourth of the aims outlined in Section 1, as the Discovery Programme's dataset expands with the delivery of data from EMIS and from the acute trusts and as the role of Discovery as a data service to support direct care locally becomes more widely known, there will be more applications from various NHS groups, academic bodies and commercial concerns to access and use the data.

Some projects have already been given approval to progress with accessing data and while progress has been slowed by issues around consistent data delivery to date, there has been the following progress against each of these usage projects:

- Atrial Fibrillation

An outline submitted in February 2017 to the Barts Charity has been approved for consideration by the Charity for funding. An outline application is being submitted in August 2017. This programme – REAL-health (PI: Robson) includes three themes:

- *Child Health (Dezateux)*
- *Respiratory (Griffiths)*
- *Cardiovascular (Timmis)*

There is funding for development of the atrial fibrillation work in the cardiovascular element of this. There is also work underway to obtain Research Capability funding for six months from 1st April 2017 to fund a GP (Luis Yerbe Garcia Morzon) for six months (1st April – 31st September 2017) to work further on atrial fibrillation linkage through the CSU.

- CEG Analysis

In full form, the project has not yet started. However, the prerequisite of overcoming technical problems of accessing the Endeavour website through the firewall has now been resolved. A number of users within the CEG are able to access the database and have started carrying out data checking activities, such as checking that numbers of (for example) females under 50 or patients on five or more drugs are the same in Discovery as they are in practice data. Where discrepancy exists, the team is checking on specific patient detail.

- Child Health

National Child Measurement Programme (NCMP) data for last three school years has now been sent by Hackney, Newham and Tower Hamlets Local Authorities under signed Data Sharing Agreements (DSA) with the CEG running until 2022. Professor Dezateux and her research assistant, Nicola Foster, are working on these datasets at CEG. Initial analyses of Hackney data were presented at recent Discovery Board. Other components of child health, namely linkage to newborn screening blood spot data, are under development in collaboration with Public Health England and Great Ormond Street Hospital NHS Trust.

- Genes & Health
The project has currently recruited 21,000 of total 100,000 volunteers and is progressing well. NHS numbers have been obtained on 18,977 volunteers and, of these, 13,770 (73%) are linked to primary care e-health records via CSU. The remaining 27% likely live outside the current four CCG regions and it is anticipated that this is something which would be addressed more effectively in Discovery. The interim Data Safe Haven solution is Aridhia.

- Natural Language Processing / Clinithink
Initial discussions have been held with Clinithink for the Natural Language Processing (NLP) work and a proposed contract drafted to cover licence and initial support costs for two identified projects within population health based themes.

Project 1 – COPD

Identify a cohort of individuals in East London who may benefit from withdrawal of corticosteroid therapy based on their social and clinical characteristics matching the set of characteristics known to be associated with beneficial corticosteroid withdrawal.

Project 2 – Prostate Cancer

Seek to identify any clinical or social characteristics that help predict which patients with prostate cancer are more likely to benefit from radiotherapy.

The current draft proposed contract covers licencing and initial support costs for both projects and the projected cost has been included in Discovery spend for 2016/17.

- Where are my patients & who is providing care?
Streamline identification of the organisations and people caring for patients from primary, community, secondary, mental health and out of hours providers, better integrating the links between the patient and all the health and social care professionals who provide care.

- Whole System Data Project
Establish an integrated health, social care and wider determinants of health dataset across Tower Hamlets which allows for effective risk stratification and resource allocation for the local population based on evidence, using Discovery East London as its data haven/processor for data to identify:

- ❖ *Socio-demographic determinants*
- ❖ *Health status*
- ❖ *Service usage.*

The project will seek to achieve sufficient granular identification of health inequalities in the Borough

Another group have been positively reviewed as first applications and are progressing to outline their use case in more detail:

- Hurley Group
This project is proposed to go live with patients extracted from EMIS into Discovery that did not fall into the original cohort of the four East London CCGs. The objective of this project is to host the Hurley Group's data within the NHS Discovery service, giving access to identifiable patient data (limited) to the patients registered to Hurley Group practices, to: -
 - *facilitate the creation of a real-time dashboard visible at all the practices with disease finding capability, e.g. number of possible diabetics; accessibility, demographics and disease burden, together with queries to help identify patients that are pre-deterioration or if a site needs support in an area that impacts health.*
 - *Interrogate the interest and interaction with the data, e.g. a GP sees that they have a large number of possible diabetics and then goes on to query why that is e.g. ethnicity. This to try and change the current trend which is "why is this an issue, can you find out for me". A simple query engine that people can drill down further into and answer the question themselves.*

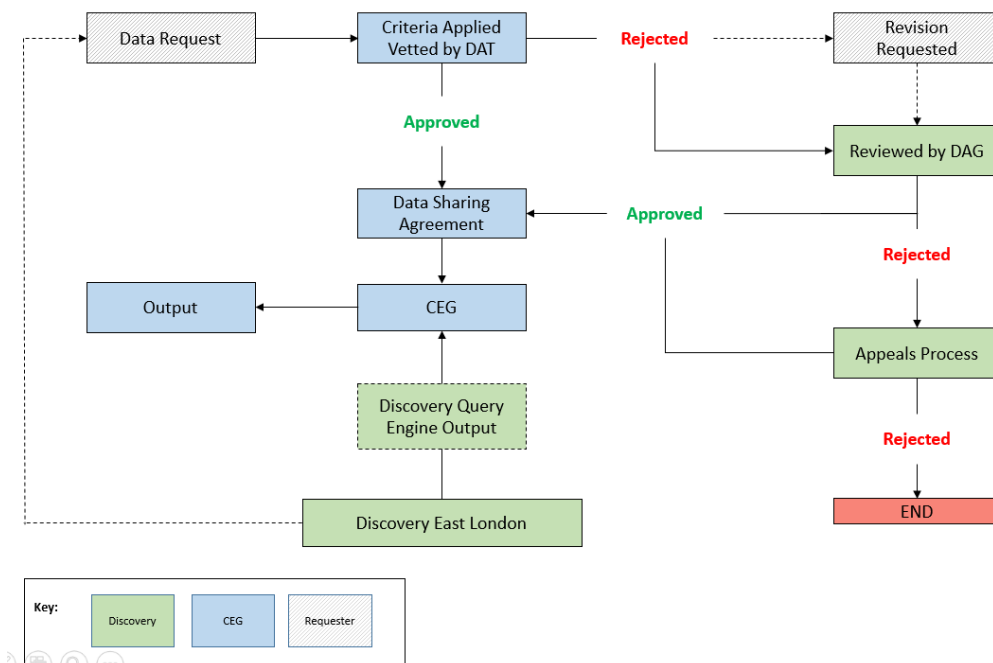
- 18 Other services who use EMIS as their clinical system
EMIS offers product to non-GP services and the 18 services below have been added to the current EMIS roll-out plan and processes are underway to facilitate data extract in May. Work will be undertaken to fully understand the implications of this:

❖ <i>0-19 Children's Services</i>	❖ <i>Older People & Rehabilitation Services</i>
❖ <i>Cardiovascular Nursing</i>	❖ <i>Primary Care Psychology</i>
❖ <i>CASH – Contraception & Sexual Health</i>	❖ <i>Rehab & Health Psychology</i>
❖ <i>Child Health</i>	❖ <i>Specialist Children's Services</i>
❖ <i>Children's Therapy</i>	❖ <i>Speech & Language Therapy</i>
❖ <i>Clinical Assessment Service (CAS)</i>	❖ <i>Wheelchair Service</i>
❖ <i>Community Dieticians</i>	❖ <i>Community Renal Service</i>
❖ <i>Diabetes Service</i>	❖ <i>Barts Health Newham Diabetes Service</i>
❖ <i>Foot Health</i>	
❖ <i>Heart Failure Nurses</i>	

There are also a range of possible or potential applications in discussion currently (as outlined in the Project Register attached as **Appendix B**) and these would seek to progress rapidly when access to the Discovery data becomes available. The Discovery team has to invest time and resources to review each application, as the numbers increase a process will be required that will allow us to prioritise those project that provide most benefits to the community we serve. A Data Access Review application template has been shaped (**Appendix C**) and is currently in use with a number of potential new applicants to access the Discovery data. The template aims to provide the Discovery East London governance groups with the detail required to assess the structure and objectives of an applying project, its relevance to the Discovery East London objectives and the extent of public benefit to be achieved by the work.

The Caldicott III review recommendations³ set the framework for a common approach to data sharing for both individual and indirect care. Current suppliers and new suppliers benefit as their tools can now operate on patient level and population level data as opposed to just the organisation that has purchased their clinical recording system. Just as there has been a move towards open source code, this marks the move towards open source algorithms that can be plugged into the health data contained in the open source code. In this way, Discovery can be open to innovation from any supplier or provider and can fulfil its aim to support research, development and planning, whether on consented identifiable data, or the depersonalised dataset, and help East London to become a research enabled community.

As outlined in Section 5, the Clinical Effectiveness Group (CEG) holds the role of the current Data Access Team (DAT) commissioned by the Discovery Board to review applications against criteria and agree access to the data as a first step. The CEG has already established elements of such an infrastructure but this needs to be tested with a range of applications to allow secure access and governance at a larger scale and in a robust form that can be replicated elsewhere and provide local assurance as a trusted organisation. The CEG’s Data Access Team is also supported by a Data Advisory Group (DAG) which will advise the Discovery Board in a more detailed manner around the specific uses of data, including scientific, data controller-specific, patient focussed and relevant to the local system around performance or commercial agreements. The current proposed model for review suggests that where there is not agreement or uncertainty about a request from the DAT, the request or a revised version of it will go to DAG for a focussed review and then, if necessary, to a central Discovery appeals process. This model is currently being refined and finalised.



Review Criteria

Both the DAT and the DAG consider applications against a set of defined criteria, shaped around the following questions:

³ <https://www.gov.uk/government/publications/review-of-data-security-consent-and-opt-outs>

- Is the data custodian and everyone who will have contact with the data involved in the project trained in information governance?
- Are all organisations involved in the project bona fide?
- Has the application demonstrated how the project will benefit patients and/or the public?
- Are there concerns about commercial interest or involvement in the proposal?
- Does the application outline in sufficient detail the engagement and communication activities with individuals contacted by the project?
- Are individuals aware of the use of their data or is the use wholly compatible with the purpose for which data was originally collected?
- Is the data requested highly sensitive or relating to a vulnerable population?
- Are the information security controls outlined in the application adequate?
- Does the application post any privacy or ethical concerns?

8. ONGOING SUSTAINABILITY

Research Bid Submissions

Advancing Applied Analytics – Health Foundation

A bid submitted in July 2017 to the Health Foundation for resource to support Discovery developing a predictive analytics system by examining the risk of developing diabetic foot and amputation in our population using data assets taken from both primary and secondary systems in East London. The data describing features of diabetes progression will be submitted to machine learning techniques to develop a scalable and generalisable solution for healthcare predictive analytics in other clinical situations. The feature analysis will be shared with data partners in Wales, Scotland and Northern Ireland to test the generalisability of the predictive tools and to test the feasibility of applying the solution to large data set

Early Diagnosis Group – Cancer Research UK (CRUK)

Lung cancer is the commonest cause of cancer death across north east London with wide variation in treatment rates and survival seen, despite attempts at case mix adjustment the underlying reasons for this remain unclear. The proposal to CRUK (to be submitted on 21st August 2017) outlines a project to use linked primary and secondary care data to investigate this variance with a view to making systematic change. This will be achieved by defining potential actionable moments and incorporating real time feedback to facilitate earlier referral with a view to reducing variance, increasing radical treatment rates and improving patient outcomes.

REAL-Health Bid

Following submission of an outline, the team has applied for the 15th August 2017 application date with a bid offering Discovery an opportunity to get new informatics skills and capacity. The team will circulate the full application to relevant Discovery team for comment before submission.

SHARE: - Secure Health Analysis and Research in East London

An application to the Barts Charity for £150k was successful and an informatician (Marian Priebe) has been recruited to the relevant post and a second applicant is being considered. These informaticians will work on:

- a. Curation of codes/concepts/searches
- b. SQL data access, linkage and extraction.

Substantive Sites – Health Data Research UK (HDR UK)

Programme attendance at the preliminary information session for this call, which expects to identify 5-7 Substantive Sites to address challenging research priorities identified by the HDR UK, which would be difficult to fund through traditional funding mechanisms. The Substantive Sites will be supported through long-term awards (10years+), with approximately £25m initially invested by HDR UK over this first five-year award. An emphasis on working in partnership is strongly encouraged. HDR UK Sites will be expected to take a leadership role, not only across local partners, but more significantly across the UK to deliver the research priorities of the Institute at scale. There is a significant opportunity for HDR UK to engage with a London-wide substantive site through a QMUL-led research initiative, among others.

STP-Wide Relationships

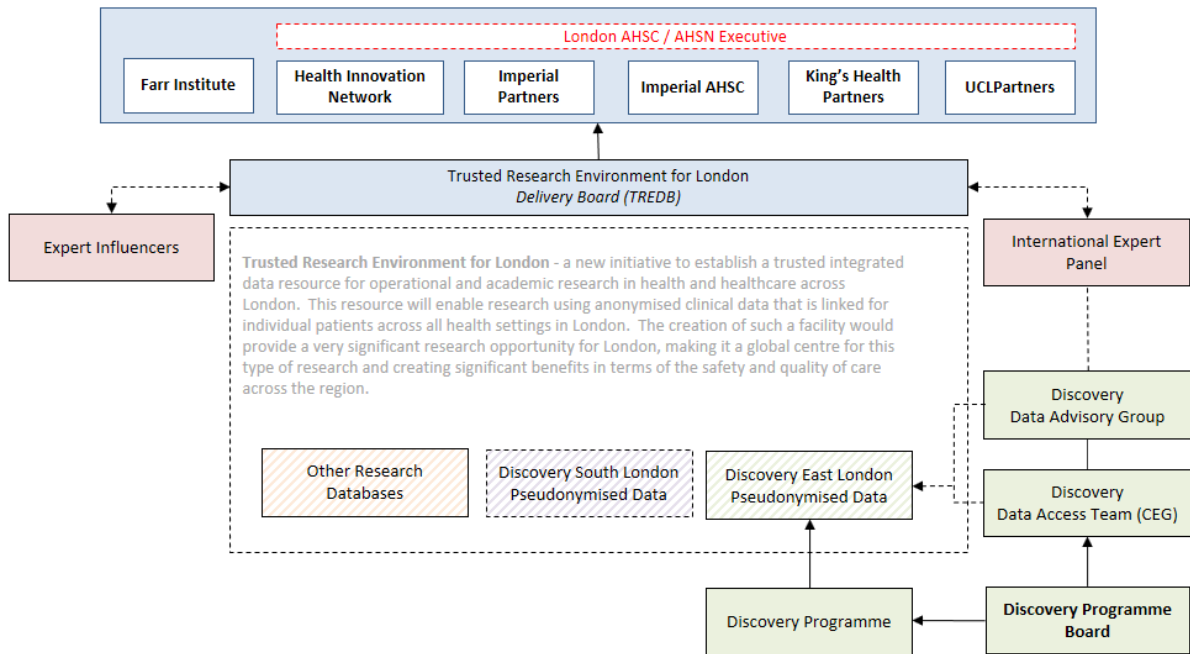
Discussions are underway with local community, mental health and social care providers across the four London Boroughs. Individual use cases are in development to test the scope and nature of the ongoing relationship between the London Boroughs and Discovery.

With a view to extending the current partnership membership to make it coterminous with the broader geography of the East London Health & Care Partnership geography, discussions are underway with teams in West Essex and Barking, Havering & Redbridge to explore opportunities for joint working and the infrastructure/system discussions required to create interoperability across the geographies. Linked with this work, initial work with Care City in Barking is underway to identify potential links between the NHS test bed for innovation and Discovery East London as part of the work on Care City as a strategic priority for UCLPartners.

London-Wide Relationships

Discovery East London is an acknowledged and linked contributor to the AHSC/AHSN Trusted Research Environment for London (as visualised below), aiming to feed linked primary/secondary care dataset into the broader London-based safe haven for research purposes.

Links with wider national Data Safe Haven networks have been made, including the European Data Safe Haven Network under the aegis of the national Farr Institute.



9. ONGOING PROJECT PLAN

A detailed Discovery Programme plan is in development in line with the specific requirements of the Discovery Programme Board and the individual applicant project plans. It will provide the detail to sit behind the programme's current logic model (**Appendix D**).

APPENDIX A:

DISCOVERY EAST LONDON Frequently Asked Questions

April 2017

1. What is the Discovery project?

New Project in East London based on a new data service set up to support direct care locally and support the development of a learning health information system that has benefits for direct quality of care and research to improve health at a population level. For robust analysis of population-level outcomes, a complete data set is often required

It covers the 4 CCGs of TH, C&H, NH and WF and the Barts Health and Homerton NHS Trusts with a population coverage of ~1.5million (200 GP practices).

Aim is to link up, subject to sharing agreements patient data to improve direct patient care. An example: if a sharing agreement is set up between GP practice and school nursing service, then GP practice would have on the EMIS clinical system, records of HPV Vaccine given by the School Nurses.

Secondary aim is to produce a depersonalised linked database similar to what CEG does with primary care data but extended now to include secondary care.

The project is focussed on combining the fractions of complete patient data held by different providers e.g. Secondary Care trusts and GPs, extracting all data including free text from every health and social care provider. As a first step, the project is focusing on data processing from local GP practices and hospitals (Barts, Homerton, Newham).

2. Who is funding the project?

The project is largely funded by the Endeavour Healthcare Charity but each of the four CCGs (TH, CH, NH, WF) has contributed to a figure of £225k shared between the organisations. The project is hosted by the CEG.

3. What kind of data is extracted?

All data is extracted including free text. Free text in consultations could be useful for different specialities e.g. in diabetes it would be important for clinicians to know why patient stopped taking a drug and the reasons could have been written in free text. Free text also allows Discovery to explore the use of Natural Language Processing to help identify patient risk from detail included in the record's free text data.

4. How is this different from HIE?

HIE simply 'views' data from various providers in the sharing network. No aggregate information is held on servers, access is governed by a data sharing agreement and HIE data cannot be used for analysis. Discovery allows combined shared data to be held on its cloud-based server and this can be analysed to improve care.

5. Do we need to get patient consent?

Patient consent is not required at this stage. As Discovery is acting as data processor for the data controllers and dealing with patient data fed into the system. The practice remains the data controller in the same way as the current systems (e.g. EMIS) arrangements. Further details about consent will be provided in later stages of the project.

6. How is data shared and held?

Data is extracted from clinical systems via direct application interfaces. The data is held on Discovery servers providing secure access via NHS standards-based APIs. Access is governed by NHS standard security model, including data sharing rules and access for users with specific roles. The individual providers (e.g. GP practices) remain the primary Data Controllers and retain control over their data within Discovery at all times.

7. Who has access to the data?

At this stage, nobody has access to the data apart from the Discovery Project Team and members of the CEG team who will be involved in validating the accuracy of the data in Discovery. Access will be determined by approved information sharing agreements between the data controllers and Discovery as data processors choose to progress.

8. How will access to the data work in practice?

The Discovery project has a Board made up of IT / informatics leads from the individual CCGs, CEG and the acute Trusts. It will also shortly include patient representatives. The CEG would approach the Discovery Board with any request to use the held data-sets for projects initiated by any of the four member CCGs, acute trusts, related academic teams or broader health and social care providers across North East London. Any request for access by a private company would also go to the Discovery Board for full review prior to approval.

Following approval by the Discovery Board, CEG-level practice data would be transferred to the Discovery servers with agreement for use by the approved applicant. This would also be in line with the access requirements set by external organisations such as NHS England or NHS Digital, for instance.

9. How does this fit in with the spine?

All patients have their demographic details on the spine via the Patient Demographic Service. Discovery is not related to the spine in any way, currently. Were this to change in the future, practices would be informed. No clinical data from Discovery passes to national systems and Discovery is not connected to the Summary Care Record.

If patients wish to opt out of the Summary Care Record (SCR), there are Read codes for GPs to enter within EMIS for this. That blocks a data flow from EMIS to the SCR. Signing up to Discovery does not affect this.

10. How will patients opt out of having their records shared with other contributors to Discovery

If patients do not want their records shared with other specific contributors through Discovery, that will be handled within the Consent Manager module of Discovery. Patients will have full control over who their records are shared with.

It will work like this:

- All patient data is replicated within the practice domain and under practice data controllership within Discovery.
- No data will flow into a specific organisation's domain unless the practice has a sharing agreement with that organisation.
- Patient dissent to this flow trumps everything, so patient dissent will block this flow or indeed any other flow from the practice domain within Discovery.

11. What control do patients have over their anonymised data?

The Data Protection Act and the recommendations of the National Data Guardian, to be published as part of the Caldicott Review, regard anonymised aggregated data as being publishable. For example, data detail at the level of: *"There are 2,053 diabetics at your practice, 27% of whom smoke"*. Discovery will reflect best practice in disclosure control; for example, as used by the Office of National Statistics to ensure patient confidentiality is protected.

12. What about pseudonymised depersonalised data?

Pseudonymised depersonalised data is data where the NHS number is replaced with pseudo ID, date of birth with age etc. If this is handled in a secure environment, this is not regarded in law as personal information. However, practices would obviously want to respect patient wishes on this use of their data, where possible. Dame Fiona Caldicott's recommendations on the use of this data has been published, consulted on and the government is about to respond. We await the final report and when the Discovery Board gets sight of the recommendations, it will decide how to best comply with this. The recommendations will outline exactly what rights patients have over their depersonalised data. For example, it is very likely that informing Cancer Registries etc. will be compulsory.

13. What do I have to do?

The CEG holds the main agreement with Endeavour and all other providers are required as co-signatories. Individual GP practices, as data controllers, and providers (including the large Trust) are being asked to sign a Data Processing Agreement (DPA). In addition to signing the Discovery DPA, practices need to activate the agreement in EMIS:

Click on EMIS web button > Configuration > Data Sharing Manage > My Agreement > Data Distribution > Developing a learning health system in the east of London 24037 > Activate My Agreement

Appendix D: Discovery East London – Logic Model (June 2017)

