FAIR TREATMENT

Federated analytics and AI research across TREs for child and Adolescent Mental Health

DARE UK - Phase 1

Dr Anna Moore, Clinical Lecturer in Child Psychiatry
Mr Tony Evans, Director Redstone Consulting

Lead: University of Cambridge

Partners: NHS Cambridgeshire and Peterborough NHS Foundation Trust; Cambridgeshire County Council; University of Essex; Eastern AHSN; Anna Freud National Centre for Children and Families; University of Birmingham; Bitfount; Kaleidoscope Data Privacy Consultants; InterMine; AIMES

Aims

Using the identification of childhood mental health problems as our use case, we will:

1. Combine existing technologies (below) to create and implement a reusable pattern for a federated trusted research environment, based on the Five Safes, that enables federation of cross-council data. Technologies are:
   a. AIMES (TRE provider)
   b. InterMINE (software enabling integration of cross council data)
   c. Bitfount (software enabling privacy-preserving federated analytics).

2. Examine the unique governance issues that arise, co-creating an aligned governance model that enables cross-council data use and its federation, which is acceptable to public, patients, and data contributors.

We will show we can answer questions such as:

What is the incidence and prevalence of adverse childhood experiences and childhood mental health problems in children 0-17y, across Cambridgeshire, Birmingham, and Essex?

Bitfount

AIMES

InterMine

Technology workstream

Progress and planned activities

Progress to date

- Cambridge TRE infrastructure designed, configured, and built by AIMES
- TRE requirements captured for demonstrator TREs (Essex and Birmingham) - these will be used to demonstrate federated analytics.
- Reviewed input data sources and designed the InterMine core data model for synthetic datasets
- Developing InterMine / Bitfount integration
- Decisions made around access controls and infrastructure updated accordingly.

Next steps

- Technical design planning & building the demonstrator TREs (Essex and Birmingham).
- Bitfount to develop tooling to import data models across multiple InterMine instances
- Demonstrate CRATE de-identification across different TREs.
- Import schemas from Welsh SAIL database (mental health) and i-THRIVE (education and social care) to create synthetic dataset
- Deploy security tooling into TRE to harden access controls.
- Document TRE Technical & Security Protocol.

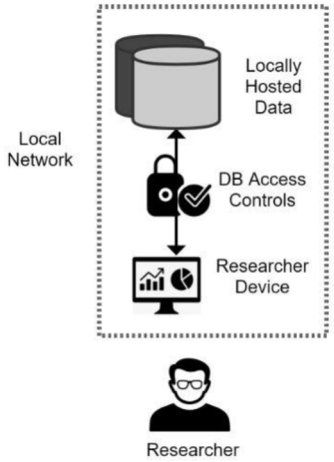
Issues & User Needs

Need a technology platform that addresses user and governance needs.

1. ***Fine grained access controls*** - Provides the tools to ensure governance and ethics compliance while supporting data re-use.
2. ***Data virtualization*** - Data collected and stored in different TREs is likely to have different structures, virtualisation enables them to be presented to researchers in a similar manner.
3. ***Data harmonisation and standardisation*** - Health, social care & education data is stored using different standards, these need to be mapped and presented to researchers in a unified manner.
4. ***Unified interface*** - Creating a single interface (API) to access the data across different TREs provides researchers with simpler access. This access also needs to span different TREs.
5. ***Solutions that have acceptable Total Cost of Ownership*** - They need to be able to be administered easily by the data controllers, with testable implementations of security controls and governance requirements following defined patterns to provide confidence in the solution.

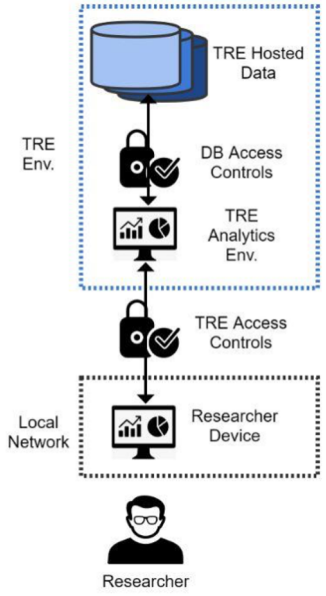
FAIR Solution Architecture

Need to understand the technology platforms that are currently supporting research and the limitations of their control models, and then how we can improve them to provide more flexibility. This follows the same evolution as from “hosting” to “managed services found elsewhere”.



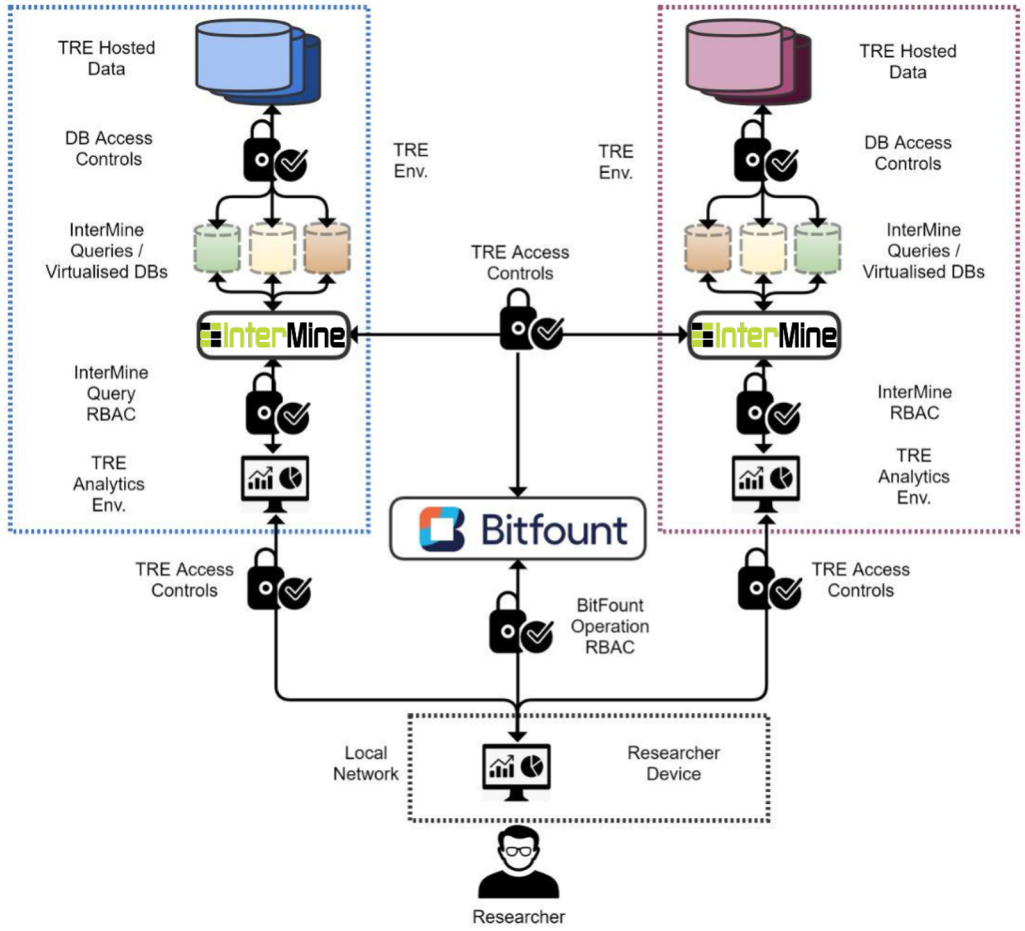
1. “Local” model

- Limited analytics capability.
- Limited IT support.
- Access controls typically on a per-database (or per-table) basis.



2. TRE-based “Infrastructure as a Service”

- Scalable analytics capability.
- IT support via Service Level Agreement.
- Access controls still typically on a per-database basis



3. FAIR-based “Platform as a Service”

- Spans TREs.
- **InterMine** support data virtualisation and interface harmonisation.
- Allows sub-database access controls.
- **Bitfount** provides ready access and allows operation-based controls.
- Tools have user-oriented interfaces, not technical interfaces.

KALEIDOSCOPE
Data Privacy Consultants

Anna Freud
National Centre for
Children and Families

Governance and PPIE programme

Progress and planned
activities

Issues

The aim is to find a governance model that can acceptably, feasibly and viably:

1. Bring together the data from health, education, social care, research
2. Enable re-identification with consent
3. Enable federated analytics
4. Utilise the full extent that the law enables
5. Acceptable to across 3 regions

- There is little appetite for novel solutions due to risk adversity
- HRA is prepared to support novelty, but local teams are not willing to ‘put their neck on the line’
- ‘Daily Mail’ anxiety

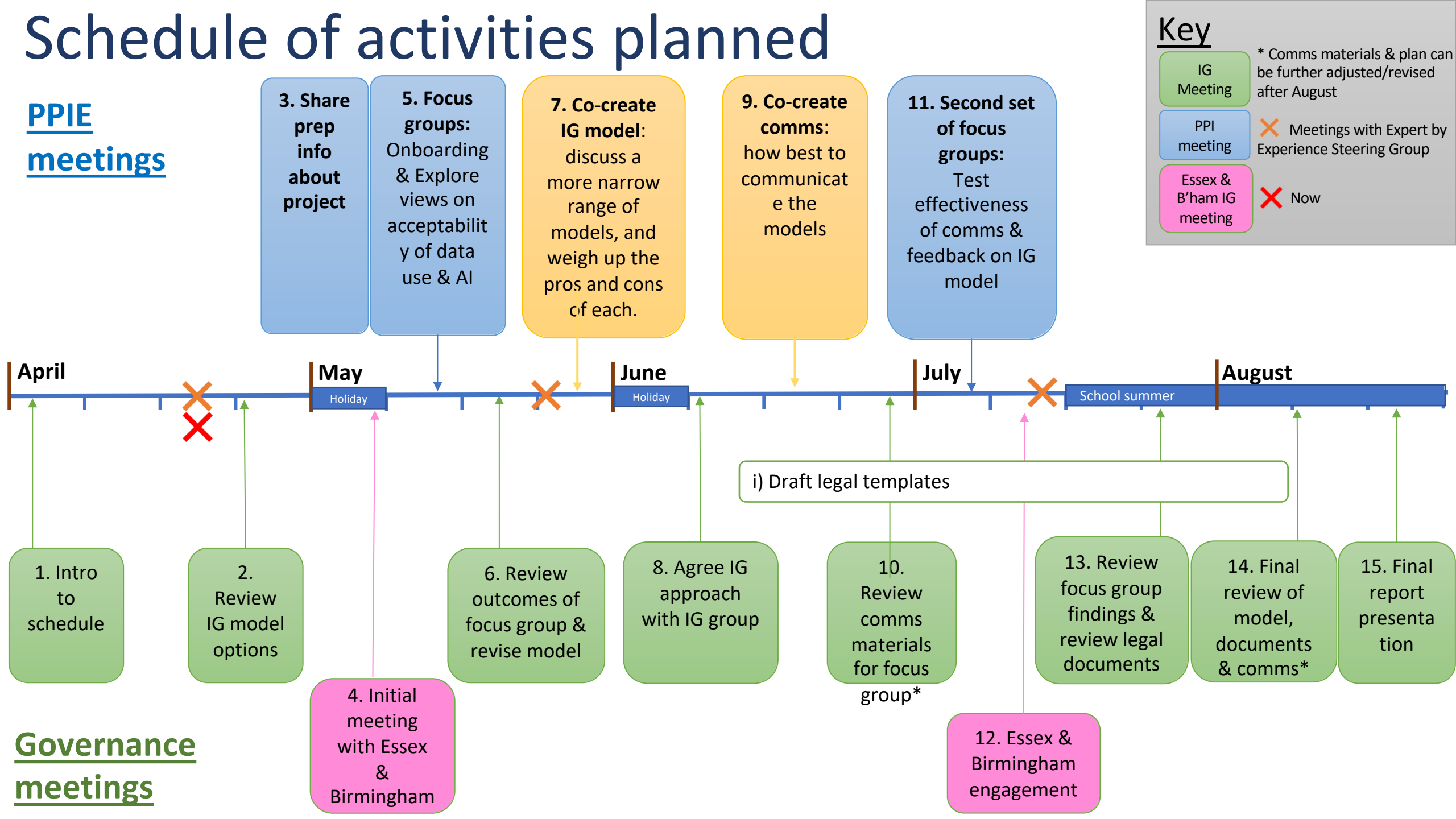
Progress to date

- Recruited PPI panel (n=142 to date, involved 40+ charities/groups to maximise diversity)
- Agreed approach to co-creation & PPI workshop structure with IG groups
- Desktop review carried out on
  - Existing legal approaches to TRE governance
  - Existing literature on public attitudes to data linkage, including any changes post Covid-19
  - This review forms the basis of Governance Model options which will be presented to the PPI group
- Relationships between technologies mapped to understand possible approaches to access controls, data flows and security measures - to feed into PPI work

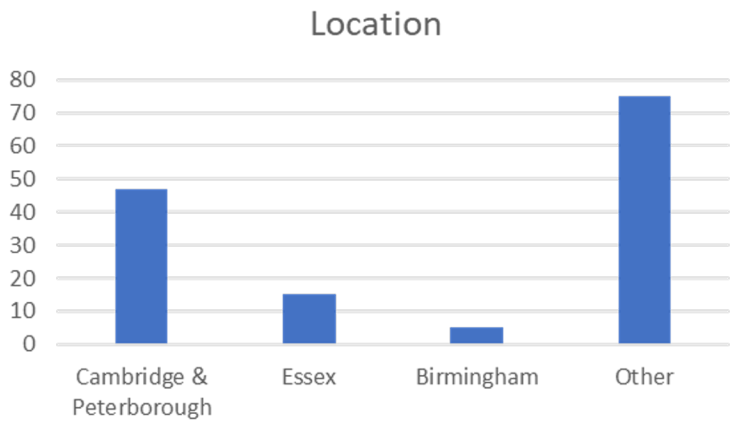
Next steps

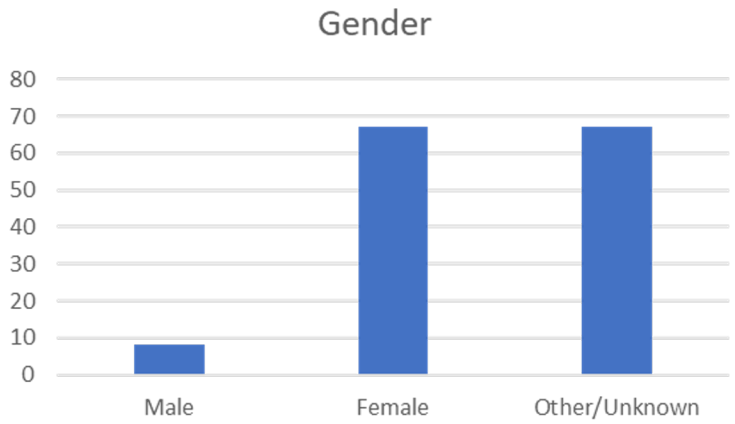
- Series of focus groups and IG groups (next slide!)
- Create PPI materials to support process
- Draft 3 potential governance models for review with IG and PPI groups
- Continue engagement with IG stakeholders in Essex and Birmingham
- Co-develop communications approach and materials

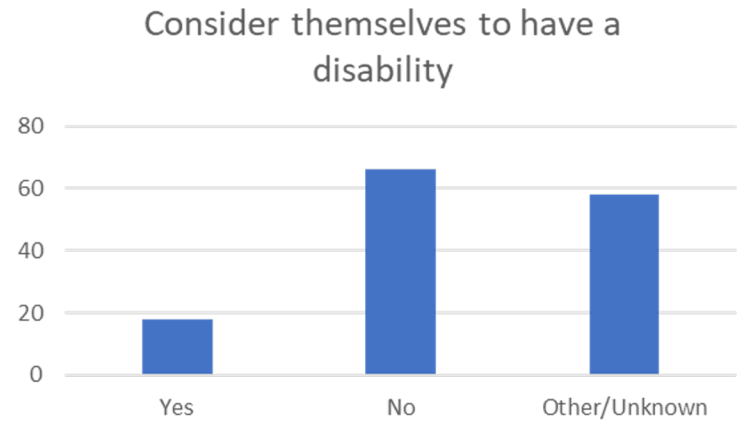
Schedule of activities planned

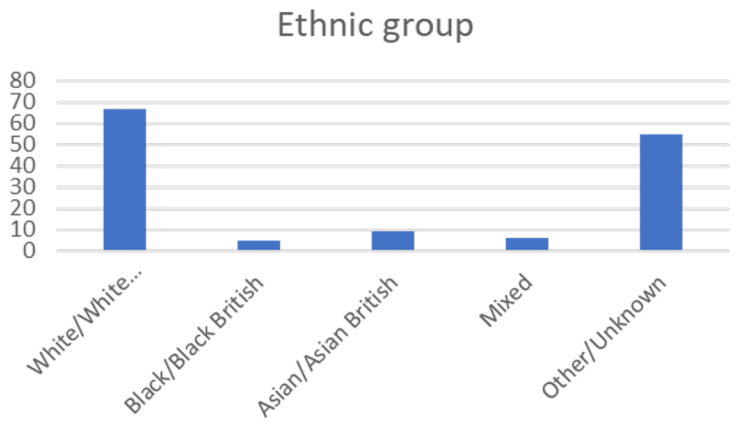


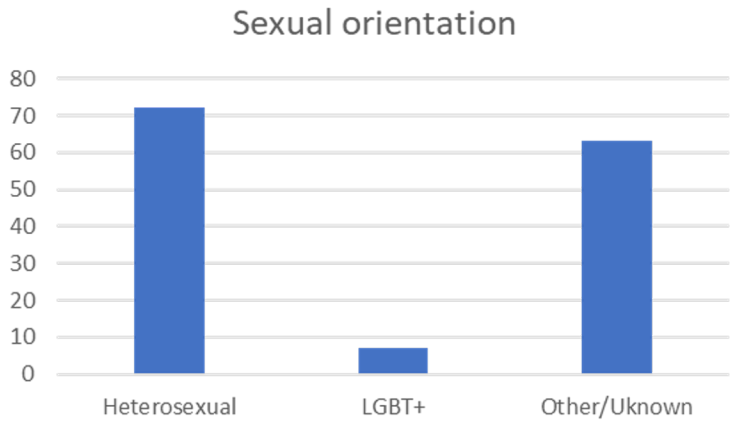
PPIE group demographics: all sign-ups (N=142)











Note: the high number of “unknown” responses reflect the fact that the EDI questionnaire was voluntary