

User survey:
<https://tinyurl.com/gel2022>



Getting medical histories for participants

Emily Perry

Research Engagement Manager

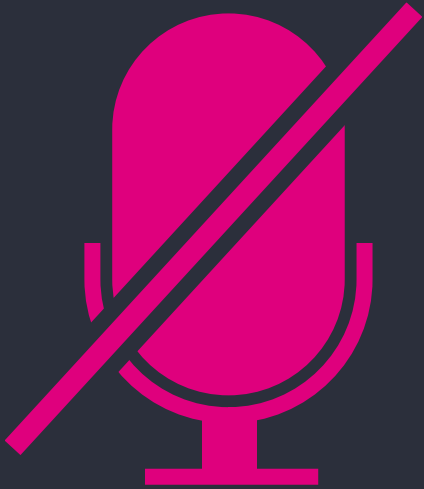
20th September 2022



Data security

- This training session will include data from the GEL Research Environment
- As part of your IG training you have agreed to not distribute these data in any way
- You are not allowed to:
 - Invite colleagues to watch this training with you
 - Take any screenshots or videos of the training
 - Share your webinar link (we will remove anyone who is here twice)
- We will record this training and distribute the censored video afterwards

Questions



Your microphones are all muted



Use the Zoom Q&A to ask questions



Upvote your favourite questions: if we are short on time we will prioritise those with the most votes.

Questions



Matthieu Vizquete-Forster
Bioinformatician -
Research Services



Helen Carter
Data Manager -
Research Data Layer



Elinor Scruby
Data Wrangler -
Research Data Layer



Michiel Scheepmaker
Senior Developer

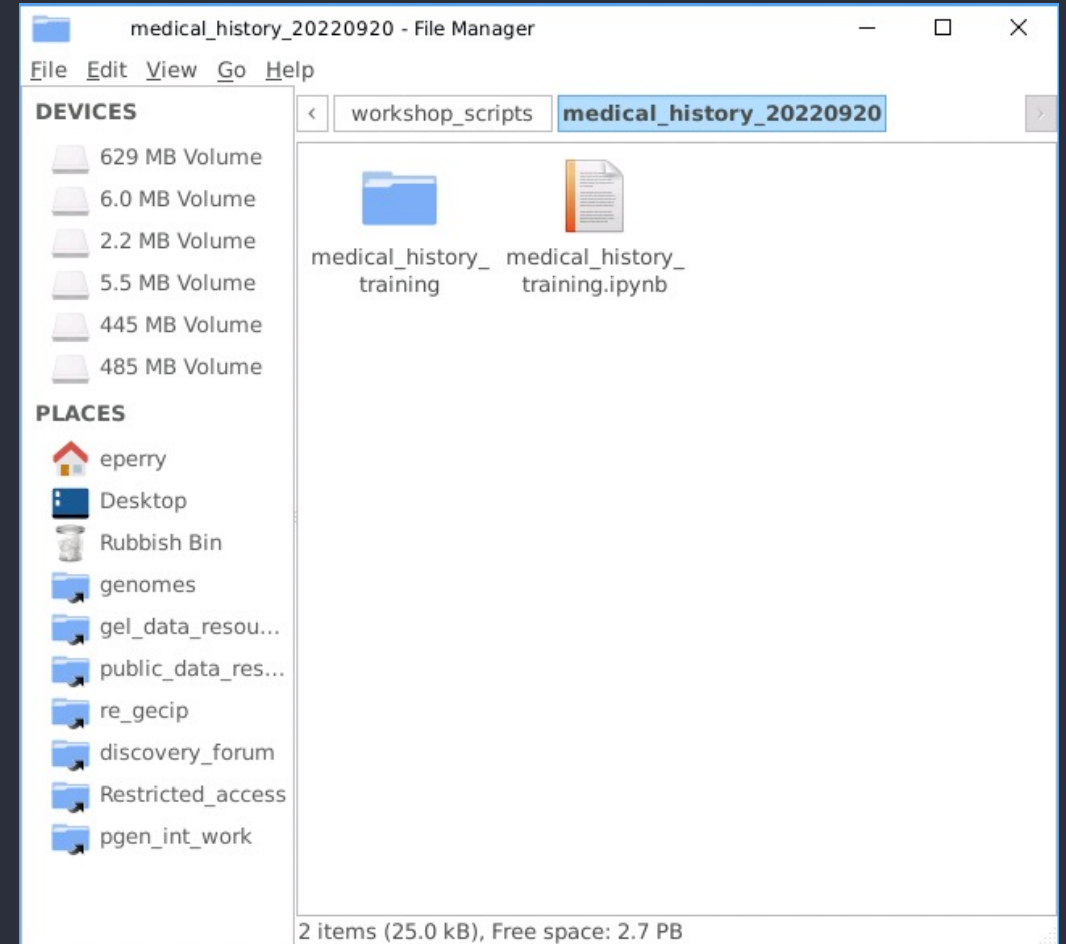
Agenda

- 1 Introduction and admin
- 2 NHS Digital data in the RE
- 3 Mental health data in the RE
- 4 Accessing NHS Digital data with Participant Explorer
- 5 Comparing participants' medical history with Participant Explorer
- 6 LabKey tables: Hospital Episode Statistics
- 7 LabKey tables: Mental Health
- 8 Accessing medical history programmatically
- 9 Questions

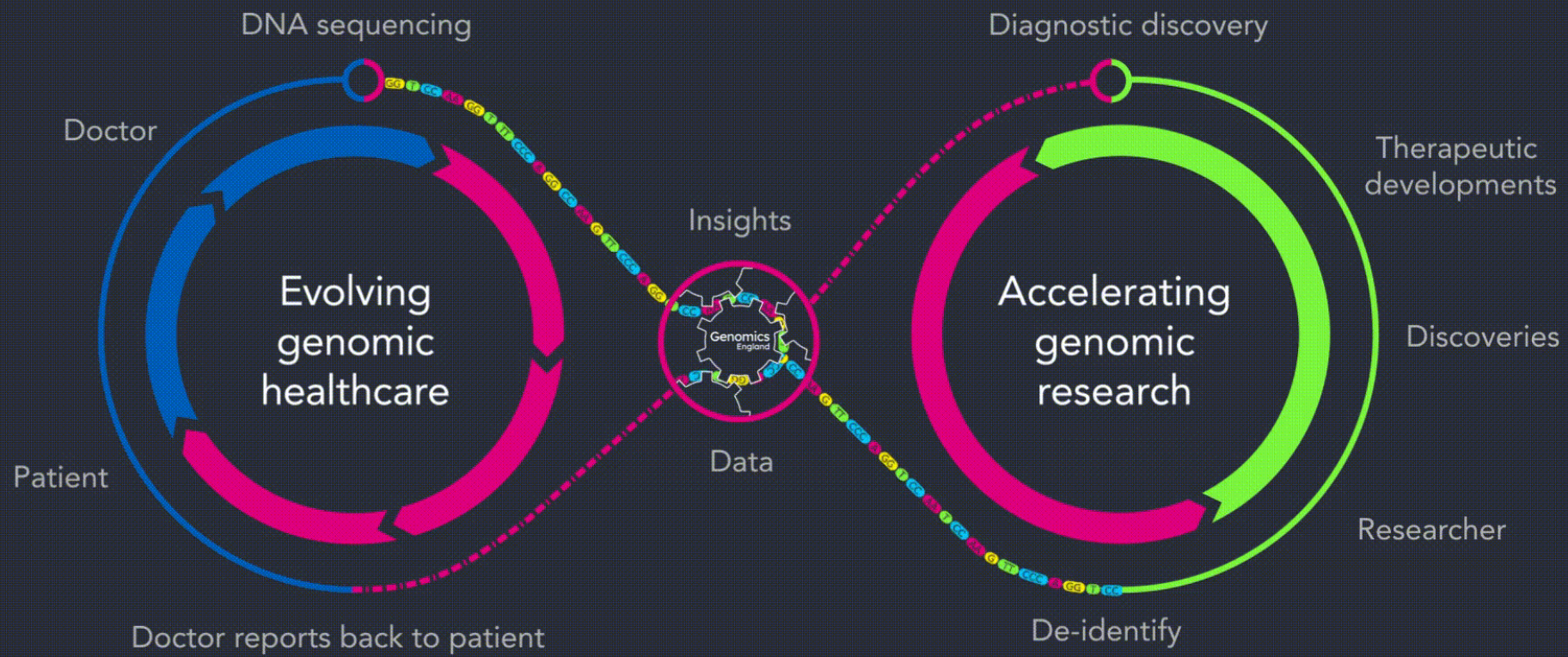


Materials

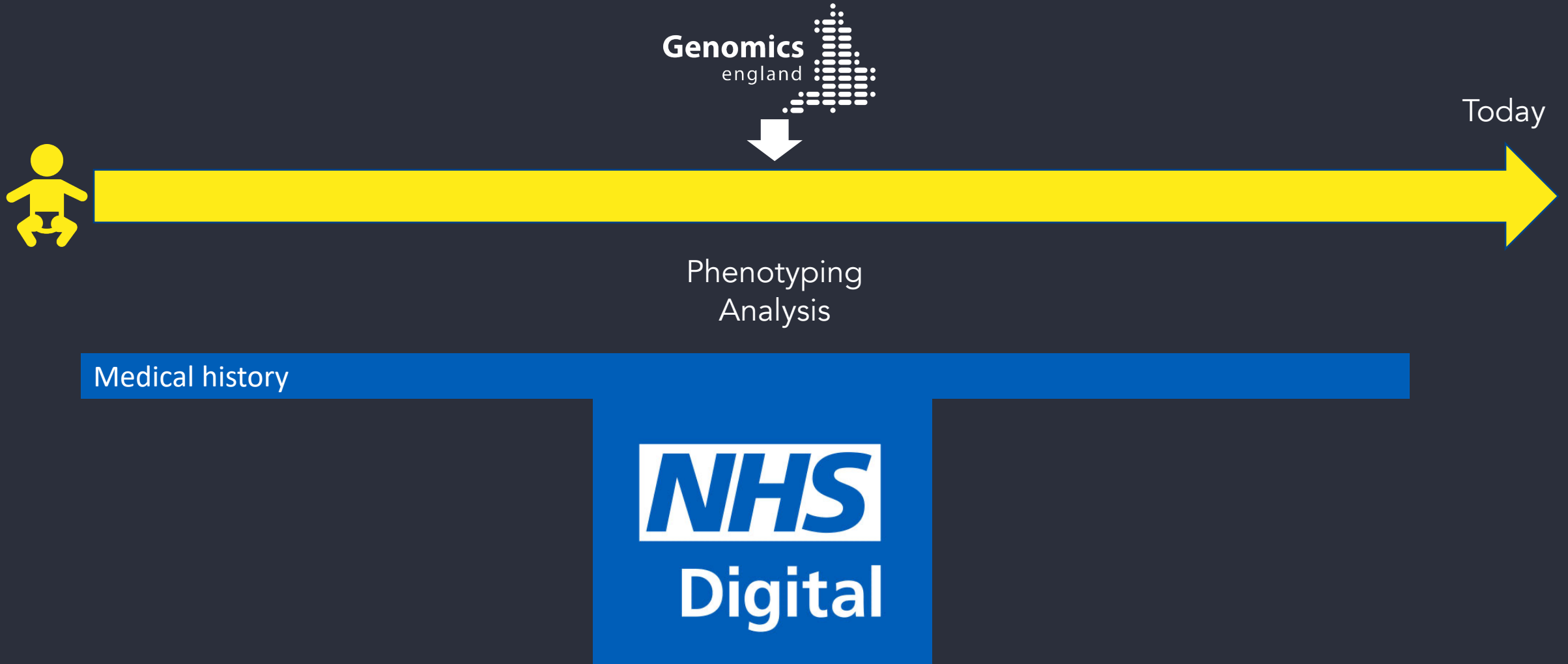
- Slides and video will be sent out to you after the session
- Scripts available in `/gel_data_resources/example_scripts/workshop_scripts/medical_history_20220920`



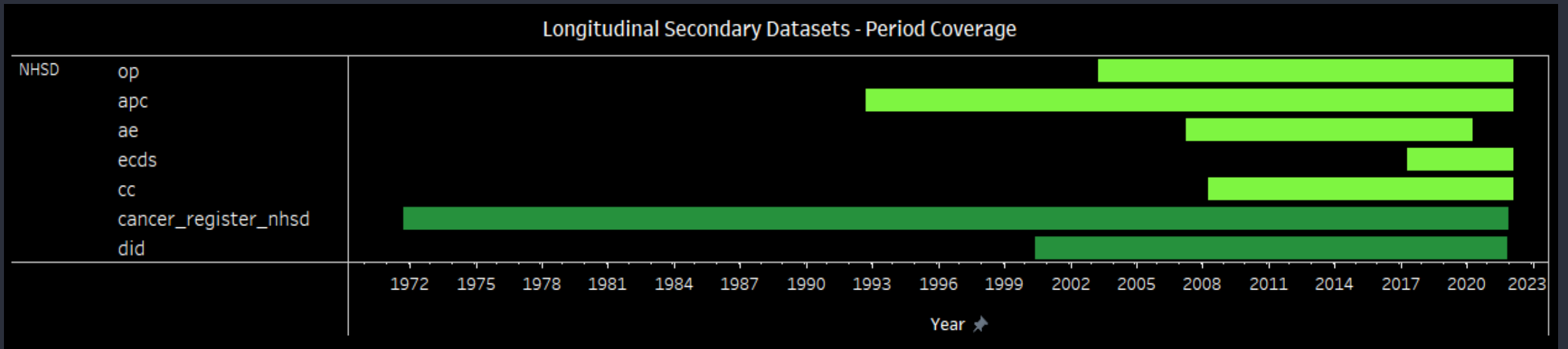
2. NHS Digital data in the RE



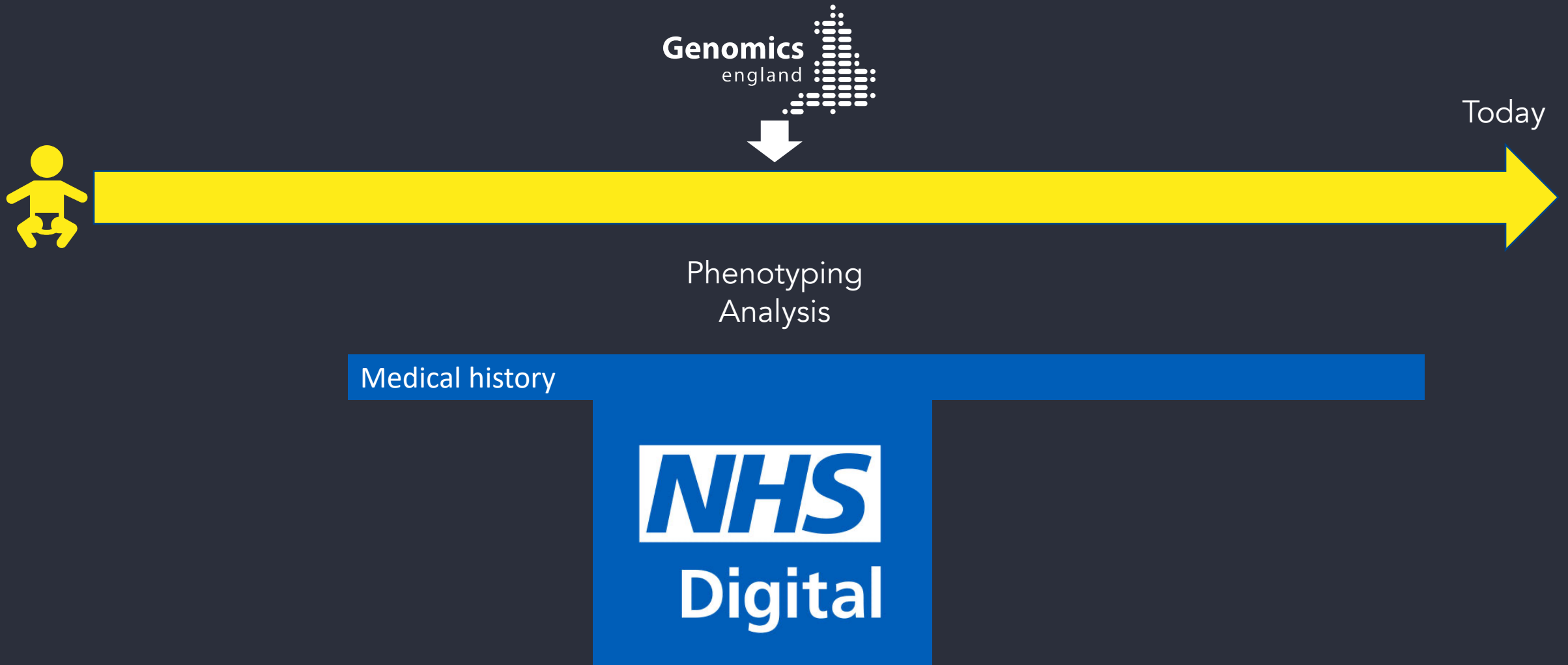
Data ingestion



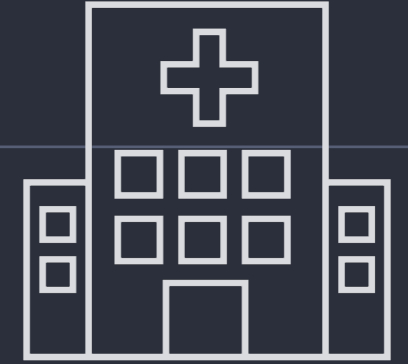
Data coverage over time



Data ingestion



Hospital episode statistics



Outpatients
op

Admitted
patient care
apc

Critical care cc

Accident and
Emergency ae

Outpatients op

Planned day appointments at a hospital

- Date/participant age
 - at appointment: `apptage`, `apptdate`
 - start/end of full series of appointments: `rttper*`
- Diagnosis codes ICD10 : `diag*`
- Procedures: `opertn*`
- Clinician speciality and type: `tretspef`, `mainspef`, `stafftyp`
- Referral circumstances: `firstatt`, `refsourc`, `refdate`
- Location/demographics/admin for NHS charging

Admitted patient care **apc**

Hospital overnight stays

- Date/age at admission/discharge
- Admission circumstances: **admimeth**, **admisourc**, **admistat**, **aekey**, **referorg**
- Discharge circumstance: **disdate**, **disdest**, **dismeth**
- Diagnosis codes ICD10: **diag***
- Treatments/procedures/investigations: **domproc**, **opertn***, **operdate***, **operstat**
- Clinician speciality and type: **mainspef**, **tretspef**
- Admission-type specific: maternity, alcohol, psychiatric
- Critical care periods:
 - Admission: **apcsour***, **apcplan***
 - Type of care: **acploc***, **apcspef***, **orgsup***
 - Discharge: **acpdisp***, **acpend***
 - Timing: **acpn***, **apcstar***, **intdays***
 - Death: **apcout***
- Location/demographics/admin for NHS charging

Critical care **cc**

Intensive care admission

- Date admission/discharge (**ccstart***, **ccdiseate**)
- Days spent under types of care: ***days**
- Admission (**ccadm***) and discharge (**ccdise***) circumstances
- Unit type: **ccunitfun**

Accident and emergency **ae**

Unplanned emergency hospital attendance – walk-ins and ambulance

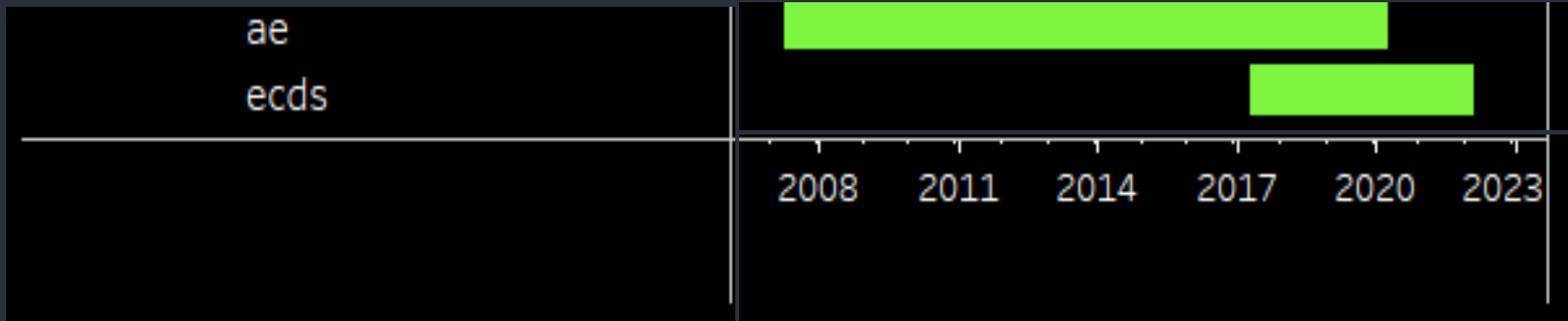
- Dates/age admission/discharge
- Diagnosis codes: ICD10 (**diag_***), type of injury (**diag2_***), anatomical area (**diaga_***, **diags_***)
- Treatments/procedures/investigations: **treat***, **invest***
- Arrival circumstances: **aearrivalmode**, **aeincloctype**
- Discharge or department/ward admitted to **epikey**, **aeattenddisp**
- Location/demographics/admin for NHS charging

Emergency care data set **ecds**

Emergency hospital care

- Dates/age admission/discharge
- Clinician speciality and type: **care_professional_tier***, **department_type**
- Diagnoses: **diagnosis_code***, **injury***, **comorbidities***
- Treatments/procedures/investigations: **investigation***, **treatment***
- Arrival circumstances: **arrivalmode**
- Discharge or department/ward admitted to: **discharge***, **referred_to_service***

ae/ecds



Hospital episode statistics

Full data dictionary:

<https://research-help.genomicsengland.co.uk/pages/viewpage.action?pageId=66847544&preview=/66847544/66847619/Data%20Dictionary%20Main%20Programme%20v15.xlsx>

Cancer registry

participant_id

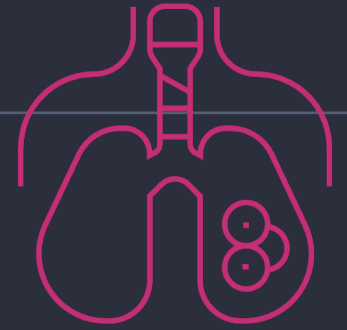
cancer_registration_year

event_date

cancer_site

cancer_type

cancer_behaviour



We don't have...



Free text



Images

did table provides list of
images taken – contact
clinician if required



GP data

Diagnostic imaging data `did`

- Dates/age: `did_date*`, `ic_age*`
- Imaging type: `did_nicip_code`, `did_snomedct_code`
- Location in body: `ic_region*`, `ic_sub_region*`, `ic_system*`, `ic_sub_sys*`, `ic_morph*`
- Related procedures: `ic_modality*`, `ic_sub_modality*`
- Source of referral: `did_patsource_code`, `did_refer_orgcode`, `ic_reftype_desc`
- Diagnosis: `ic_cancer_desc`

How do I see the images?

3. Mental health data in the RE

Mental health data

mhmd (Mental Health Minimum Dataset) – 2011-2014

mhldds (Mental Health Learning Disabilities Dataset) – 2014-2016

mhsds (Mental Health Services Dataset) – 2016 onwards



mhmd – 2011-2014

- mhmd_v4_episode
- mhmd_v4_event
- mhmd_v4_record

mhdds – 2014-2016

- mhdds_episode
- mhdds_event
- mhdds_record

record

One record for each period of mental illness

- Details of the participant
- Counts of appointments, days in hospital etc during
- Dates of starts/ends
- Scores for physical and mental ability

episode

One record for each episode of care (eg hospital stay)

- Start and end dates
- Discharge details
- Admission details
- Care speciality

event

One record for each appointment

- Date
- How, where and who with
- Assessments carried out
- Scores:
 - HoNOS
 - PHQ
 - SAC

mhsds – 2016 onwards

- mhsds_bridge
- mhsds_master_patient_index
- mhsds_gp_practice_registration
- mhsds_patient_indicators
- mhsds_care_coordinator
- mhsds_care_plan_type
- mhsds_crisis_plan
- mhsds_care_plan_agreement
- mhsds_service_or_team_referral
- mhsds_service_or_team_type_referred_to
- mhsds_other_reason_for_referral
- mhsds_referral_to_treatment
- mhsds_onward_referral
- mhsds_discharge_plan_agreement
- mhsds_care_contact
- mhsds_care_activity
- mhsds_other_in_attendance
- mhsds_indirect_activity
- mhsds_responsible_clinician_assignment
- mhsds_hospital_provider_spell
- mhsds_ward_stay
- mhsds_assigned_care_professional
- mhsds_delayed_discharge
- mhsds_hospital_provider_spell_commissioner
- mhsds_medical_history_previous_diagnosis
- mhsds_provisional_diagnosis
- mhsds_primary_diagnosis
- mhsds_secondary_diagnosis
- mhsds_coded_scored_assessment_referral
- mhsds_coded_scored_assessment_act
- mhsds_coded_scored_assessment_cont
- mhsds_care_programme_approach_care_episode
- mhsds_care_programme_approach_review
- mhsds_clustering_tool_assessment
- mhsds_coded_score_assessment_clustering_tool
- mhsds_care_cluster

mhsds – 2016 onwards

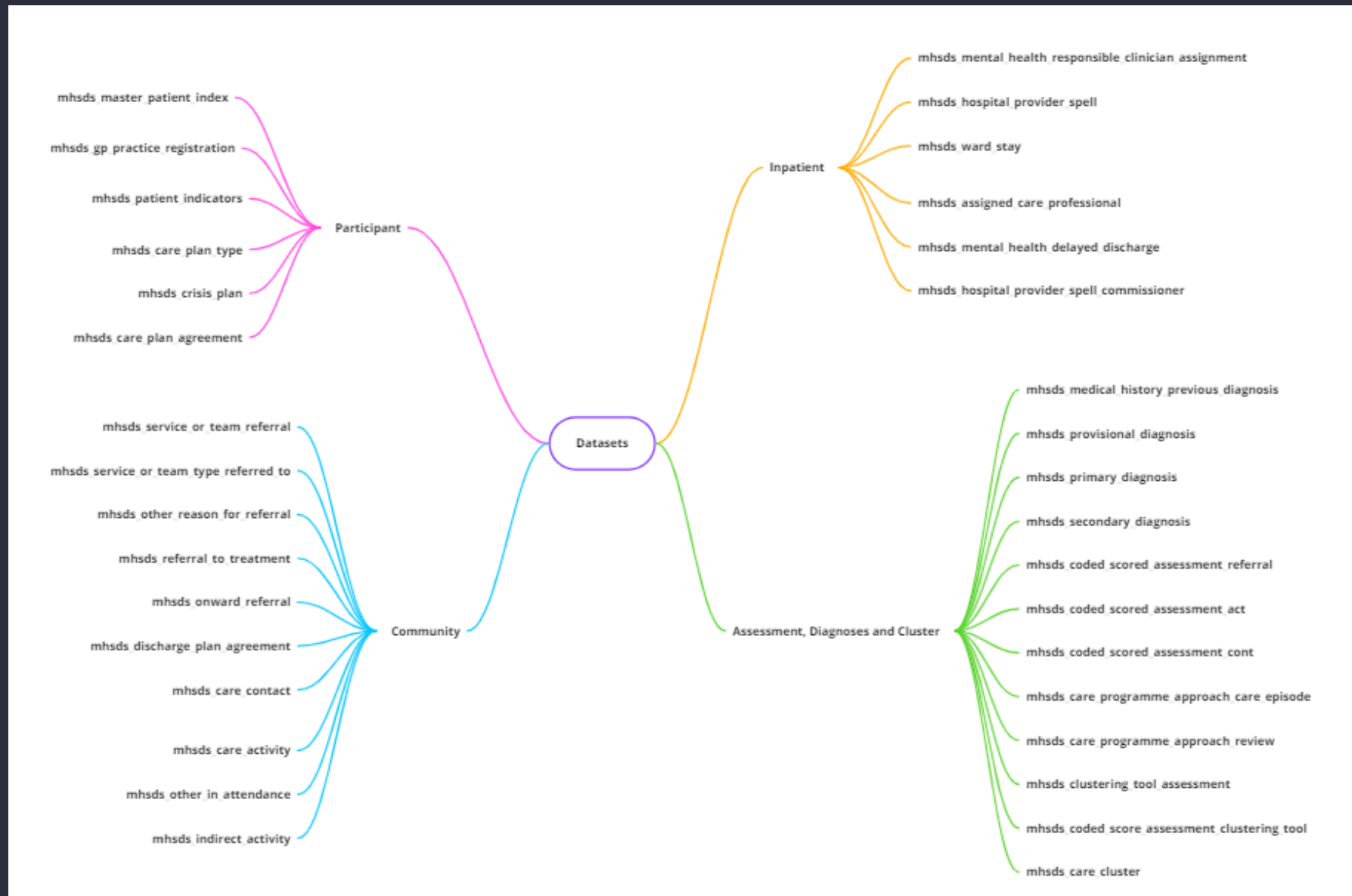
mhsds_curated_participant Overview of general participant information; demographics, death details, GP registrations and psychosis indicators, as well as details of any care plans created for a participant.

mhsds_curated_community Overview of community (outpatient) care. This includes details on referrals, discharge agreements and care contacts with associated care activities.

mhsds_curated_inpatient Overview of inpatient care. This includes details of hospital spells, ward stays, delayed discharge periods and associated clinicians and care professionals.

mhsds_curated_assessment_diagnoses_and_cluster Overview of scored assessments and clustering tool assessments completed, patient diagnoses and allocated care clusters.

mhsds – 2016 onwards



mhsds_curated_participant

- Participant details: date of birth, sex, ethnicity
- Death details
- Care plan details
- Psychosis and antipsychosis medication start

mhsds_curated_community

- Discharge details: `dis*`
- Referral details: `primreasonreferralmh, ref*`
- Care activities: `carecont*`
- Observations and treatment codes: `*code*`
- Dates of appointments/meetings

mhsds_curated_inpatient

- Dates/times: estimated/actual
- Type of care: `hospitalbedtypemh`, `intendclincareintencodemh`, `specialisedmhservicecode`, `ward*`
- Treatment: `treatfunccodemh`

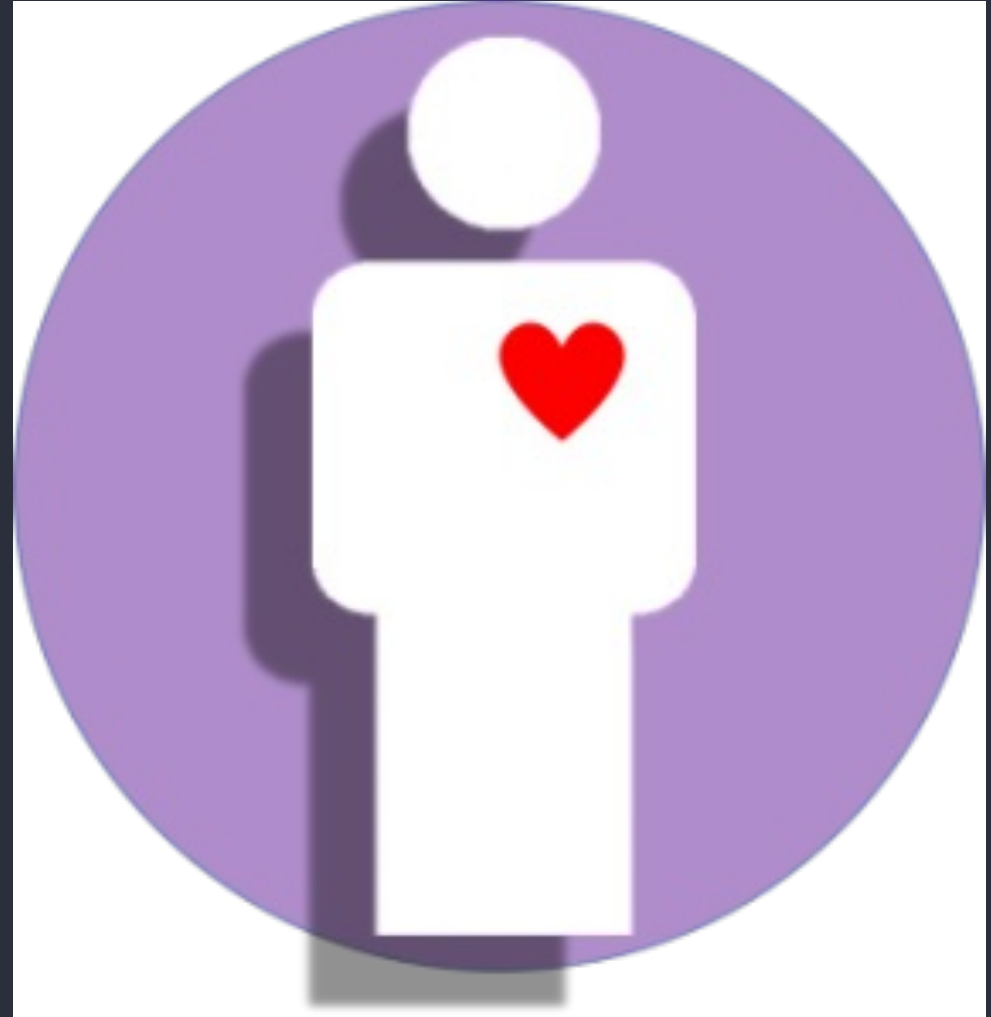
mhsds_curated_assessment_diagnoses_and_cluster

- Assessment tools and scores: `csa*`
- Clustering tools: `clust*`, `*clustinit`, `*clustcodefin`

4. Accessing NHS Digital data with Participant Explorer

Participant Explorer

- Point-and-click tool
- Ontology-aware
- Natural language
- Tables of medical encounters
- Graphical display
- Compare participants



5. Comparing participants' medical history with Participant Explorer

6. LabKey tables: Hospital Episode Statistics

7. LabKey tables: Mental Health

8. Accessing medical history programmatically

LabKey API

Labkey API allows you to:

- Combine data and filters from multiple tables
- Work in a variety of programming languages, but most support for Python and R
- Work both locally and on the HPC

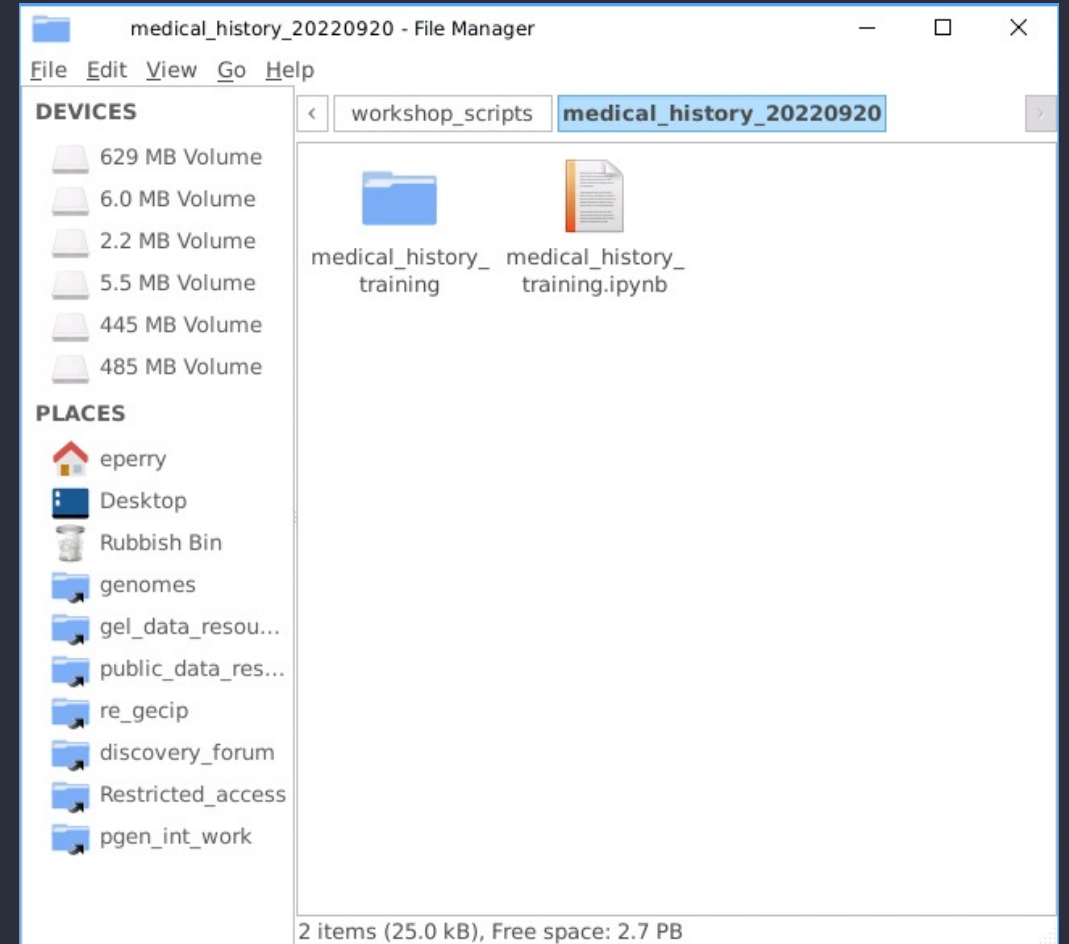


Set up .netrc

- You can access the same data via the LabKey API as you can through other means
- You will need to configure access to the LabKey API with your username and password
 - In your home directory
 - On the HPC
- You do this by editing a file called .netrc

Materials

- Slides and video will be sent out to you after the session
- Scripts available in `/gel_data_resources/example_scripts/workshop_scripts/medical_history_20220920`



Accessing the notebooks

Python

```
module load python/3.8.1  
jupyter notebook --port  
<four digit port number>
```

Open link in browser

R

```
module load R/4.0.2  
rstudio
```

Key takeaways

Medical history
from NHSD

Easy access
with Participant
Explorer

Detailed tables
in LabKey

9. Getting help and questions

Getting help



Check our documentation:

- <https://research-help.genomicsengland.co.uk/>
- Click on the documentation icon in the environment



Contact our Service Desk:

- ge-servicedesk@genomicsengland.co.uk

Questions



Your microphones are all muted



Use the Zoom Q&A to ask questions



Upvote your favourite questions: if we are short on time we will prioritise those with the most votes

Future sessions

Using the HPC to run jobs

15 Nov

Future sessions – 2023

Jan	Introduction to the Research Environment
Feb	New datasets in the RE: GMS, Covid and pilot data
March	Using the Research Environment for clinical analysis
April	Importing tools and data for using in the Research Environment
May	Building cancer cohorts and survival analysis
June	Building rare disease cohorts with matching controls
July	Finding participants based on genotypes
August	Getting medical records for participants
September	Tools created by researchers - a showcase
October	Using GEL data for publications and reports
November	What tools and workflows should I use to fulfil an overall goal?
December	Running workflows on the HPC and Cloud

Past training

In the User Guide:

- Redacted slides and videos
- Q&A from sessions

Sessions:

- Introduction the Research Environment, March
- Building a cohort based on phenotypes, May
- **Finding participants based on genotypes, July**

Meetup this evening

Time: 6pm

Place: Revolution Cambridge

Registration: <https://www.meetup.com/cambridge-genomics-meetup-group/events/287384685/>

Feedback

User-survey on login page

Thank you