



Data for R&D PROGRAMME

The future of OMOP
across the NHS Research
SDE Network

Dr Phil Waywell

Our R&D Mission

Delivering
rapid access to
the world's
largest linked
health datasets



Improve Care



Support Innovation



Sustain the NHS



Six use cases

Data for R&D investments are guided by six key, high level use cases



AI/algorithm development
(testing, training, and validation)



Clinical trial activities
(feasibility, recruitment, efficacy through short- and long-term trial follow up)



Real world studies (safety, effectiveness, cost effectiveness)



Translational research (academic discovery and implementation of discovery into practice)



Epidemiological studies (large cohorts for population health research)



Health systems research
(evaluation of systems or processes, including operational and applied research)

A decorative graphic consisting of a grid of small grey dots that tapers to the right, positioned in the upper left quadrant of the slide.A decorative graphic consisting of three parallel diagonal bars in shades of green and blue, extending from the bottom left towards the top right, positioned in the lower left quadrant of the slide.

NHS Research SDE Network

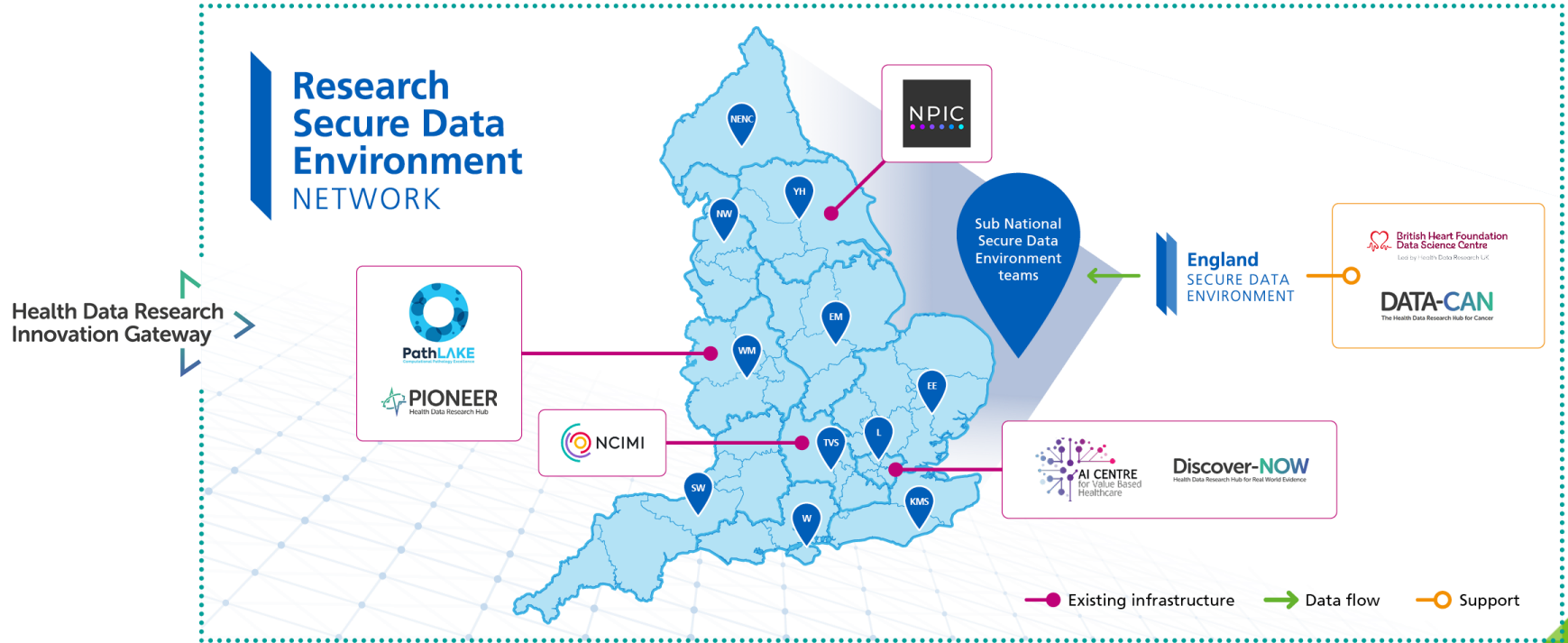
We are addressing needs through a federated Network of SDEs



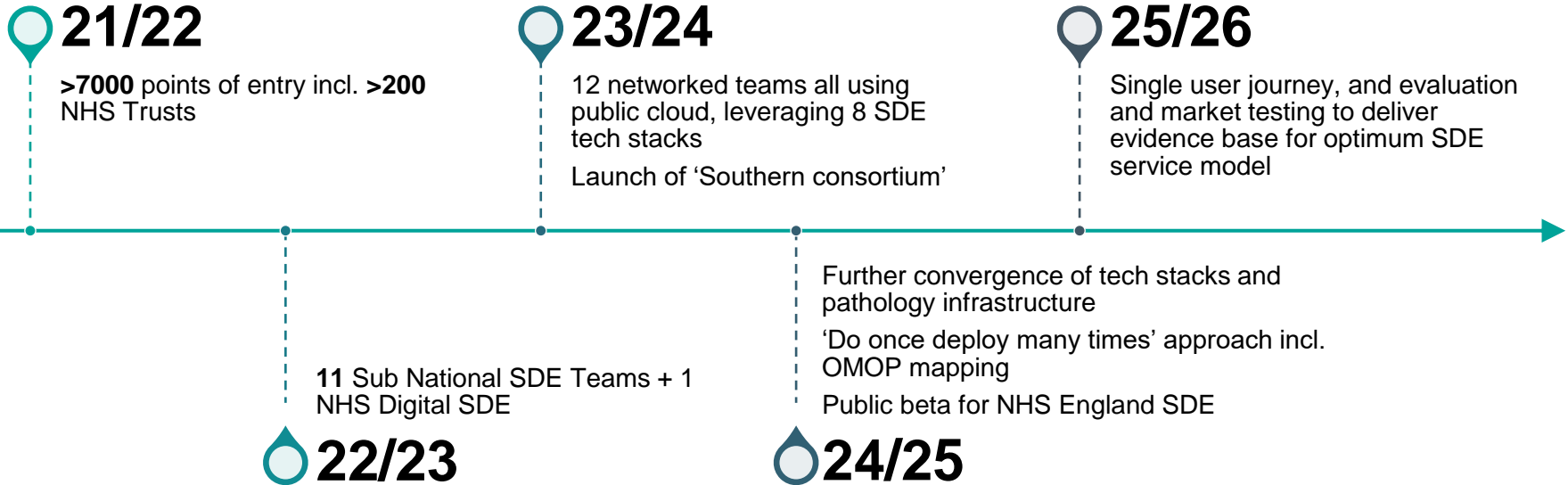
- A single programme delivering access to 57 million population scale, whilst nurturing regional innovation
- Full coverage of rich multimodal data across imaging, pathology, genomics and structured data
- Access to NHS and academic partnerships for delivery for expertise and translational connectivity
- Underpinned by patient and public involvement and engagement
- Leverages full ecosystem capacity and prior investments



Leveraging prior investment

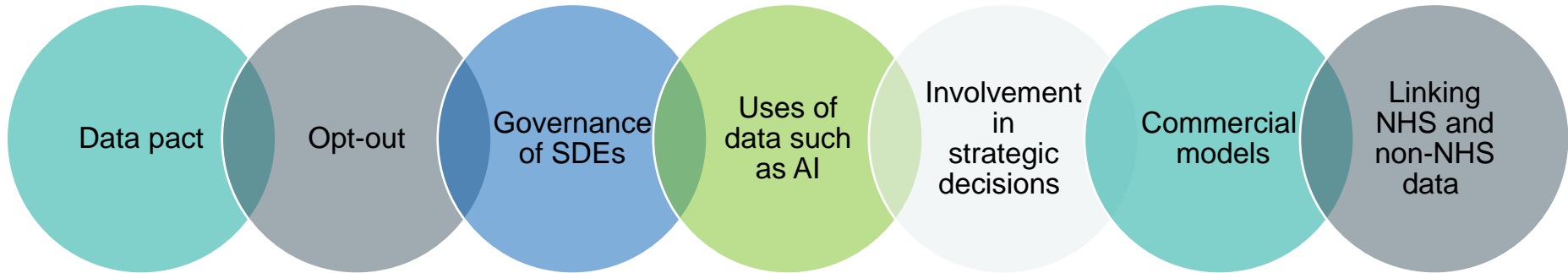


The journey so far



#PoweredByNHSDData, shaped by the public

- We have started an **extensive public engagement programme**
- We are engaging and deliberating on a range of key subject areas



- The outputs will allow the public to shape national policy, will challenge our thinking and approaches and will build public trust





Using OMOP to enable work at pace and scale

Network-wide cohort discovery

cdmTableName	cdmFieldName	rationale
CONDITION_OCCURRENCE	condition_occurrence_id	Unique identifier for each condition occurrence, facilitating tracking and analysis of individual diagnoses.
CONDITION_OCCURRENCE	person_id	Key
CONDITION_OCCURRENCE	condition_concept_id	Identifies the specific medical condition for cohort definition, disease progression analysis, and treatment research.
CONDITION_OCCURRENCE	condition_start_date	Tracks the date of condition onset for understanding disease trajectories and identifying cohorts based on diagnosis timing.
DEATH	person_id	Key
DEATH	death_date	Records the date of death for mortality-related cohorts and survival analysis.
DEATH	death_type_concept_id	Identifies the type of death (e.g., natural, accidental) for cohort definition and research on specific death causes.
DEATH	cause_concept_id	Identifies the underlying cause of death for mortality analysis and research on disease burden and risk factors.
DRUG_EXPOSURE	drug_exposure_id	Unique identifier for each drug exposure, allowing tracking and analysis of individual medication regimens.
DRUG_EXPOSURE	person_id	Key
DRUG_EXPOSURE	drug_concept_id	Identifies the specific drug for cohort definition, drug-drug interaction analysis, and medication effectiveness research.
DRUG_EXPOSURE	drug_exposure_start_date	Tracks the date of drug exposure initiation for understanding treatment patterns and defining exposure-based cohorts.
MEASUREMENT	measurement_id	Unique identifier for each measurement, facilitating tracking and analysis of individual measurements.
MEASUREMENT	person_id	Key
MEASUREMENT	measurement_concept_id	Identifies the specific type of measurement (e.g., blood pressure, lab test) for cohort definition and analysis of relevant clinical characteristics.



Network-wide cohort discovery

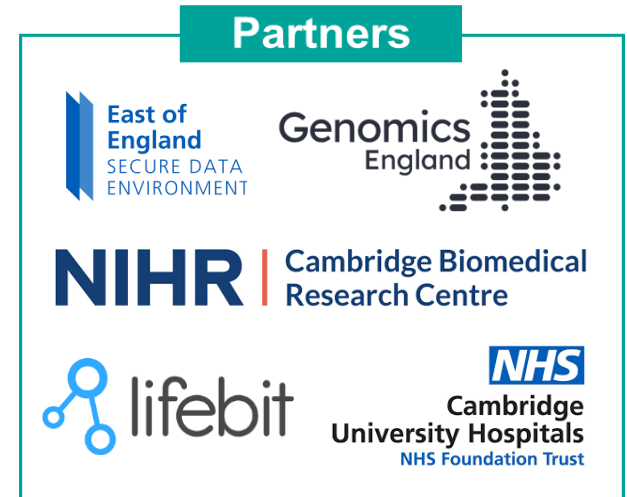
cdmTableName	cdmFieldName	rationale
MEASUREMENT	measurement_date	Tracks the date of each measurement for understanding trends over time and defining cohorts based on measurement timing.
MEASUREMENT	value_as_number	Value of the measurement
MEASUREMENT	value_as_concept_id	Value of the measurement
OBSERVATION	observation_id	Unique identifier for each observation, enabling tracking and analysis of individual observations.
OBSERVATION	person_id	Key
OBSERVATION	observation_concept_id	Identifies the specific type of observation (e.g., physical exam finding, imaging result) for cohort definition and analysis of relevant clinical features.
OBSERVATION	observation_date	Tracks the date of each observation for understanding progress over time and defining cohorts based on observation timing.
PERSON	person_id	Unique identifier for each individual, enabling linking across all tables.
PERSON	gender_concept_id	Identifies sex category for potential inclusion in sex-specific cohorts.
PERSON	year_of_birth	Determines age groups for cohort definitions and facilitates age-related analyses.
PERSON	race_concept_id	Potentially informative for understanding health disparities
PROCEDURE_OCCURRENCE	procedure_occurrence_id	Unique identifier for each procedure, enabling tracking and analysis of individual procedures.
PROCEDURE_OCCURRENCE	person_id	Key
PROCEDURE_OCCURRENCE	procedure_concept_id	Identifies the specific type of procedure for cohort definition, analysis of procedure utilisation and effectiveness research.
PROCEDURE_OCCURRENCE	procedure_date	Tracks the date of each procedure for understanding temporal trends and defining cohorts based on procedure timing.
VISIT_OCCURRENCE	visit_occurrence_id	Unique identifier for each visit, allowing differentiation and analysis of individual encounters.
VISIT_OCCURRENCE	person_id	Key
VISIT_OCCURRENCE	visit_concept_id	Identifies the type of visit (e.g., inpatient, outpatient) for cohort definition and resource utilisation analysis.
VISIT_OCCURRENCE	visit_start_date	Tracks the timing of each visit for temporal analysis and cohort inclusion based on visit timing.



Bytes, Bench, Bedside genomics driver project

Builds upon DARE UK sprint exemplar – Establishing infrastructure for secure analysis across different clinical-genomic datasets

- Automated extraction of data from host EPR systems (Epic).
- Automated the transformation of extracts to OMOP using omop_es (UCLH).
- Surfacing structured clinical data alongside genomic data for federated analysis.
- Automated release of results via an airlock.



Cancer research - expand, simplify and accelerate research into causes and treatments

Data held by the National Disease Registration Service mapped to the Common Data Model (OMOP) adopted by the SDE Network.

- 300,000 cancer cases are registered annually but data is fragmented
- Mapping the data into a common model to expand, simplify and accelerate research into causes and treatments
- Future links with other datasets in the SDE Network will enable use of analytic tools and AI to look at relationships between cancer and other conditions or lifestyle factors



SDE Network Driver Projects

Network Driver Projects serve two main purposes:

- They exemplify the SDE Network's current capability and capacity to deliver R&D studies at pace and at scales spanning multiple members of the network;
- They enable the testing and evaluation of solutions and processes required to deliver increased volumes of these large-scale studies in the future.



A decorative graphic consisting of a grid of small, light grey dots that tapers to the right, positioned in the upper left quadrant of the slide.

**Join the SDE Network
community**

A decorative graphic consisting of three parallel diagonal bars in shades of green and blue, extending from the bottom left towards the top right, overlapping the dark blue background.

Health Data User Group

- **Membership** spans industry, academia, medical research charities, NHS
- **Forum** for discussion, feedback and co-development
- **Quarterly newsletter**
- **Regular engagement programme:**
 - **Quarterly Webinars:** our next webinar is scheduled for 11:00 – 12:00, 11 December 2024
 - **Annual Conference:** 8 October 2024 in Central London.

Sign up by emailing: data.healthresearch@nhs.net

Access our FutureNHS platform: <https://future.nhs.uk/>



**Data for R&D
PROGRAMME**

Health Data User Group

What is the Health Data User Group (HDUG)?

It's a forum for stakeholders from **industry, technology, medical research charities** and **academia** who are likely to be users of NHS data infrastructure for research purposes. It provides an opportunity to **share information, hear** from other group members, and **join us** at our regular online and face-to-face events.

- The group connects the programme with health data users...
- Helping to create the right data ecosystem...
- enabling life-saving research and innovation to flourish.