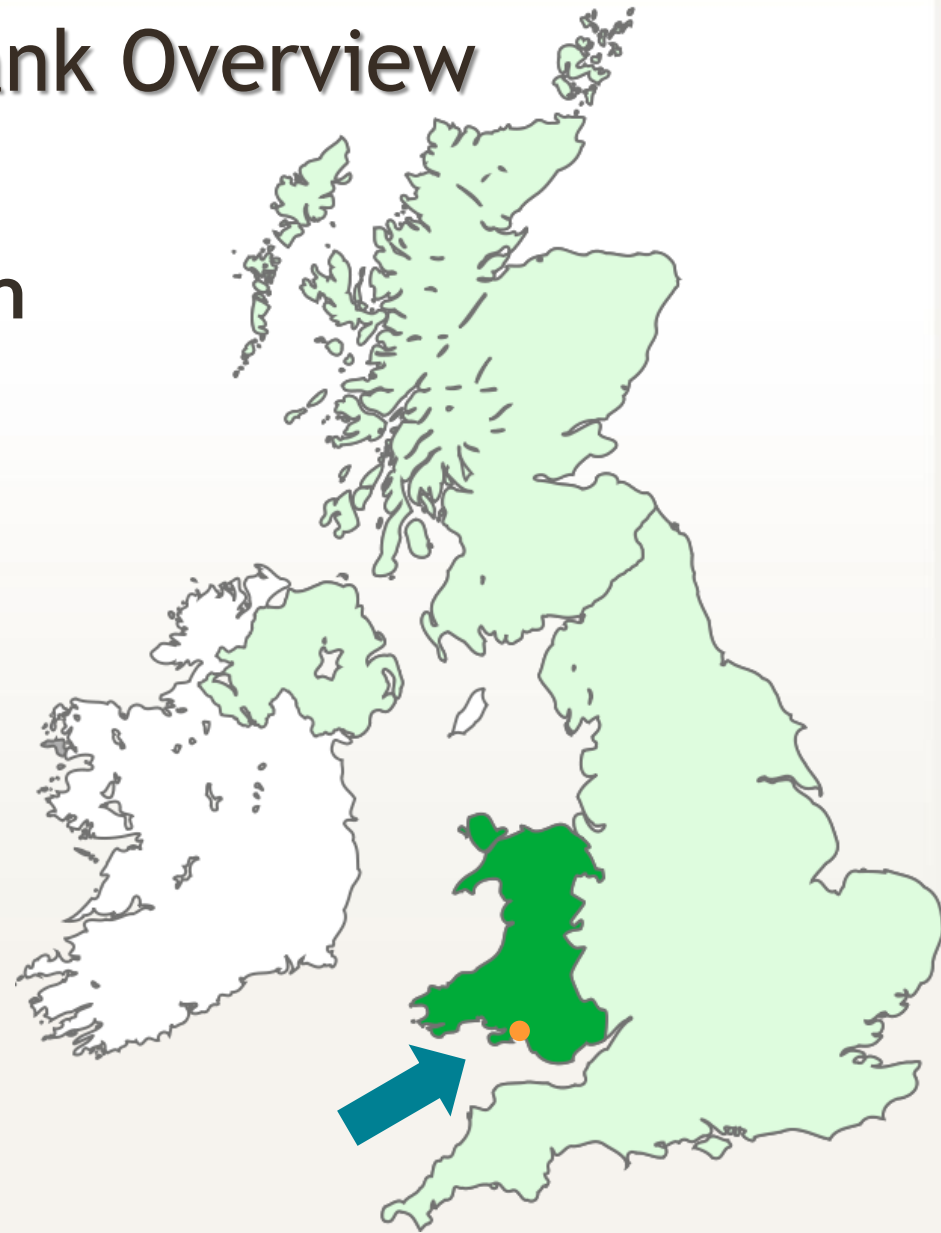# Using The SAIL Databank for RCTs

Ian Farr, Senior Research Analyst
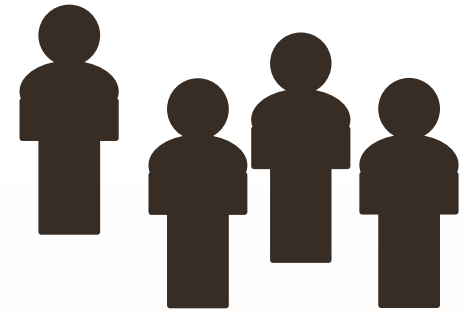
i.w.farr@swansea.ac.uk

SAIL Databank Overview

- **SAIL = Secure Anonymised Information Linkage**
- **Established 2006**
- **Swansea University**
- **Work closely with NHS**
- **Datasets:**
  - Health
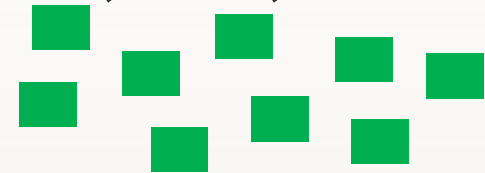 - Demographics
  - Social
  - Disease specific

SAIL Databank Overview

Welsh population-based data

Pop. 3,000,000

5,000,000

10,000,000,000

UK-Wide Collaborations

- **Health Data Research UK**
- **Administrative Data Research Network**

Some Key Linkable Datasets Held by SAIL

Demographic

ONS
Births

Deaths

Deprivation Indices

Welsh Demographic Service

National Community Child Health Database

Health

Inpatient Admissions

Outpatient Attendances

GP

Emergency Dept

Social

Education

Disease Specific

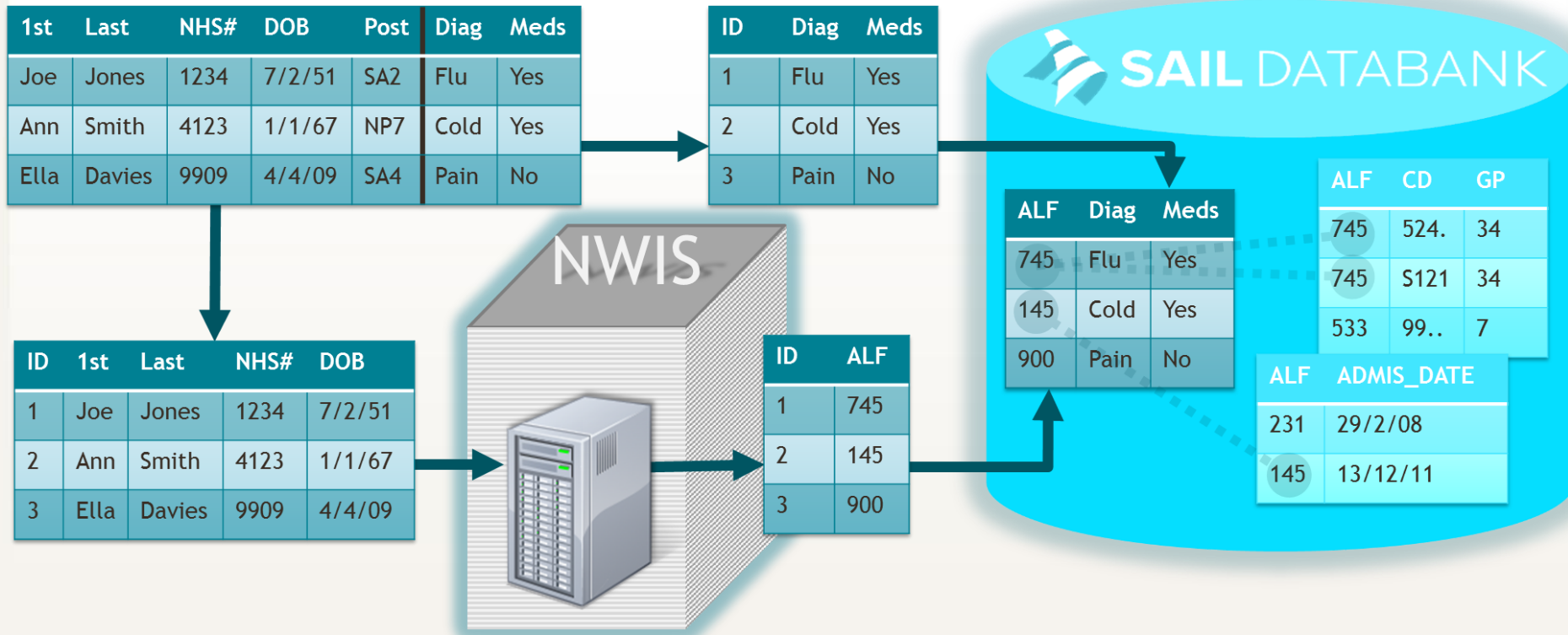Disease Registers

Screening Services

Data Coverage

- Population Coverage
  - Most datasets: Welsh population coverage
  - GP dataset: partial
    - Practices are data owners, so each must sign up.
    - In SAIL currently: 77% practices covered = 79% population coverage
    - Optional cleaning algorithms to identify record and patient sets e.g. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0228545
- Time span
  - Inpatient: 1999- (with ICD10 coding)
  - Outpatient: 2003-
  - GP: varies by practice (plenty of data 2000-)
  - A&E: 2009-
- Updates
  - Most core datasets quarterly.
  - A few datasets have a time lag at source.

Governance

- **Anonymisation of data by trusted third party.**
    - Where no NHS number then probability matching applied to derive Anonymised Linking Field (ALF) and there is a field so it is apparent this has been done
- **Independent Information Governance Review Panel (IGRP) considers all applications to use data**
    - Single application to access all SAIL datasets.
- **Projects given access to encrypted extract of the data.**
- **Work carried out in secure remote desktop environment.**
    - Review of all outputs before release.

Split file process of anonymisation

The Journey of a SAIL Project

Scoping

- Feasibility
- Refine Idea
- Determine Cost of SAIL Support

Information Governance Review

- Independent Panel
- Appropriateness of topic
- Sensitive issues

Data Loading and Preparation

- Anonymisation and loading of project datasets (if any)
- Creation of data extract

Project Active

- Project team gets access
- Conduct Analysis
- Export Results

SAIL DATABANK

Institute of Life Science
College of Medicine

Swansea University
Prifysgol Abertawe

Who Can Work with SAIL?

- A wide variety of users
  - Academic
  - NHS
  - Government
  - Commercial
- Must be legitimate research with potential for public benefit.
  - Cannot be used for performance management (e.g. league tables of hospital performance).
- Pharmaceutical companies cannot access data directly; work must be done by SAIL or another independent partner.

Types of Research: Observational Studies

- Purely based on data held in SAIL.
  - EUROmediCAT: the safety of medication in pregnancy
  - Identifying risk factors for Suicide and Self-Harm Contributes to https://adolescentmentalhealth.uk/
- Importing a cohort of patients and linking to SAIL.
  - Type 1 diabetes and hospital admissions, linking a diabetes register to SAIL.
- Deriving code lists validated from disease registers
  - E.g. https://www.hdruk.ac.uk/resources/external-validation-of-the-electronic-frailty-index-using-the-population-of-wales-within-the-secure-anonymised-information-linkage-databank/

SAIL DATABANK

Institute of Life Science
College of Medicine

Swansea University
Prifysgol Abertawe

Types of Research: Service Evaluation

- Air AWARE
  - Did an air quality monitoring and alert system in Port Talbot lead to better health outcomes?

- Housing programmes
  - Do programmes such as a home energy efficiency intervention improve health?

Types of Research: Supporting Health Providers

- Goal: work more closely with NHS to increase the benefit of SAIL.

- Small Business Research Initiative with ABMU

- GP Engagement Project
  - Research supporting GPs in focus areas.
  - Custom reports sent to each practice showing results related to their own practice.

Types of Research: Clinical Trials

- Supplemental Data
  - Replacing survey data for things like participants' health, health economics, service usage.

- Extended Outcomes
  - PROBAT, study of probiotic supplements in infancy.
    - Extra years of follow-up using routine data.

- Trial Feasibility
  - Potential for easily determining whether patients exist for a clinical trial.

SAIL Support and Costs

- No charge for data or basic service
- Project-specific support costs only
  - Different Levels of Support
    - Providing the data only
    - Collaborating on analysis
- No charge if the Information Governance Review Panel application is not successful
- Based on three types of projects
  - Academic
  - NHS
  - Commercial

Thank You

Any questions?

More info: www.saildatabank.com

saildatabank@swansea.ac.uk