



# Generation Scotland Cohort Profile: An update on the longitudinal family health study for genetic research

Hannah Milbourn<sup>2</sup>, Archie Campbell<sup>1,2</sup>, Robin Flaig<sup>1,2</sup>, Daniel McCartney<sup>2</sup>, Anne Richmond<sup>2</sup>, Heather Whalley<sup>2,3</sup>, Cathie Sudlow<sup>1,2,4</sup>  
1. Usher Institute, University of Edinburgh  
2. Centre for Genomics and Experimental Medicine, University of Edinburgh  
3. Department of Psychiatry, University of Edinburgh  
4. Centre for Clinical Brain Sciences, University of Edinburgh



## Generation Scotland - Background

Generation Scotland (GS) is a **family-based cohort study** established as a resource for research into the **genetic and environmental** determinants of **physical and mental health**. Recruitment took place from **2006-2011**.

# 24,084 Participants

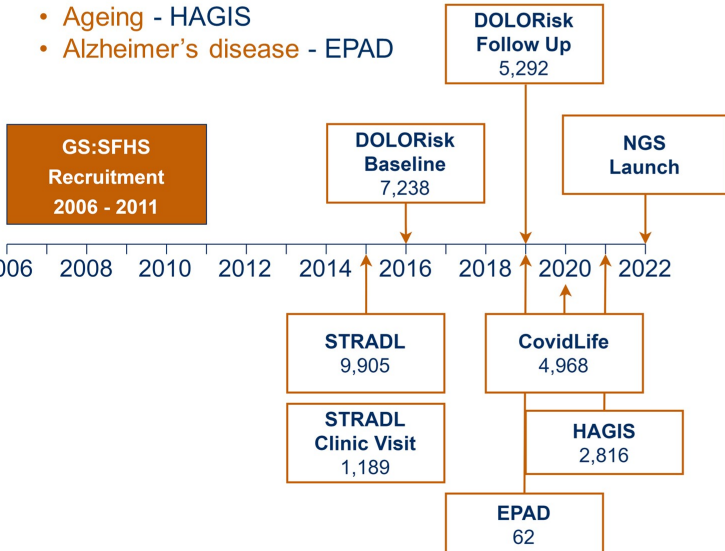
# 5,501 Families

- **Baseline questionnaire** 24,084
- **Clinic visit** 21,476 (89%)
- **NHS data linkage** 22,402 (93%)
- **Genotype array data** 20,026 (83%)
- **DNA methylation data** 18,869 (78%)

## Re-Contact Studies

Participants consented to re-contact for **prospective studies**, which has led to several additional **data collections** relating to:

- **Neuropathic pain** - DOLORisk
- **Depression** - STRADL
- **COVID-19** - COVIDLife
- **Ageing** - HAGIS
- **Alzheimer's disease** - EPAD



## NGS: New Recruitment

- In 2019 funding was obtained to expand the existing GS cohort. **Next Generation Scotland** (NGS) aims to:
- Recruit **20,000 new participants** (with expanded eligibility to **12-17 year olds**).
  - Collect saliva samples by post for **genotyping**.
  - Use established methods for **data linkage** to NHS data.

## Data Linkage

Participants consented to linkage to **NHS medical records** and other relevant sources which offer **longitudinal phenotype information**. Linkage is conducted using the Community Health Index (CHI) number.

- 147K** SMR01 Hospital admissions (N=19,611)
- 4.7M** Dispensed prescriptions (N=21,838)
- 13.5M** GP records (N=19,675)
- 4.4K** Scottish Cancer Registry records (N=3,606)
- 1.5K** Diabetes Registry records (N=1,423)

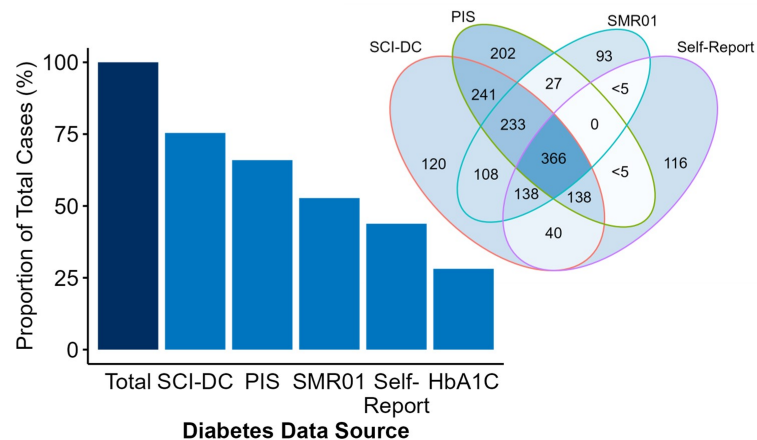
## Cohort Morbidities

Longitudinal **disease incidence** was obtained through health record linkage from **primary** and **secondary care** sources. Number of cases of common diseases in the GS cohort from **1980-2022**:

- 2,517** Hypertension
- 2,484** Depression
- 2,154** Asthma
- 1,582** Heart Disease

As an example, **diabetes cases** can be captured in **five data sources**:

- **SCI-DC**: Scottish Care Information-Diabetes Collaboration
- **PIS**: Prescribing Information System
- **SMR01**: General/Acute and Inpatient Day Case dataset
- Laboratory testing data for **HbA1c**
- **Self-reported** baseline questionnaire



## Data Access

Research proposals are subject to review by the **GS Access Committee**. Details of the application process and conditions for access are available at the study website. <https://www.ed.ac.uk/generation-scotland/for-researchers>

### Acknowledgements

Generation Scotland received core support from the Chief Scientist Office of the Scottish Government Health Directorates [GZD/16/6] and the Scottish Funding Council [HR03006]. Genotyping of the GS:SFHS samples was carried out by the Genetics Core Laboratory at the Wellcome Trust Clinical Research Facility, Edinburgh, Scotland and was funded by the Medical Research Council UK and the Wellcome Trust (Wellcome Trust Strategic Award "Stratifying Resilience and Depression Longitudinally" (STRADL), Reference 104039/Z/14/Z).

