

Ethnically Diverse Biobanks as a Foundation for Equitable Precision Healthcare

Gadaleta E*, Zhu Q*, Dayem Ullah AZM, Kolvekar T, Thorn GJ, Jones LJ, Ottewell PD, Speirs V, Kocher H, BCNB, PCRFTB, Chelala C

Corresponding author: c.chelala@qmul.ac.uk

Lay Summary

Precision medicine relies on research data that reflect population diversity, yet ethnic minority groups remain under-represented in many large UK research resources. We compared **UK Biobank**, **100,000 Genomes Project** (100k GP) and two biobanks focused on breast cancer and pancreatic cancer collections — **Breast Cancer Now Biobank** (BCNB) and **Pancreatic Cancer Research Fund Tissue Bank** (PCRFTB) — to assess representation, engagement and research use.

Large national resources provide scale and statistical power, while disease-focused biobanks offer greater cancer-specific granularity and improved ethnic diversity, particularly in geographically diverse regions such as North East London. **Embedding diversity at the point of sample collection strengthens inclusive research and equitable precision healthcare.**

1 Introduction

Precision medicine aims to tailor prevention, diagnosis and treatment by integrating genetic, environmental and lifestyle information. However, its success depends on biospecimens and data that reflect population diversity. Under-representation of ethnic minority groups in foundational research resources introduces bias into discovery, risk prediction and clinical guidelines, limiting equitable clinical impact.

Biobanks are central to precision medicine and increasingly link longitudinal biospecimens with clinical records, imaging and multi-omic data. While large national resources provide scale and statistical power, disease-focused biobanks may offer greater cancer-specific granularity and improved representation through geographically targeted recruitment.

Ensuring equitable representation in biobanks is essential, as they operate at the earliest stages of the research lifecycle and shape downstream study design, trial enrolment and translational outputs.

2 Aims

Aim. To evaluate diversity, participant engagement and translational potential across major UK biobanking and genomics research infrastructures.

Objectives.

- **Compare** demographic representation across UK Biobank, 100k GP, BCNB and PCRFTB.
- **Assess** tumour characteristics, age at presentation and outcomes across ethnic groups.
- **Examine** consent patterns, data completeness and research utilisation.
- **Evaluate** how disease-focused biobanks complement national-scale resources.
- **Identify** opportunities to enhance inclusivity at the earliest stages of translational research.

3 Preliminary Findings

A. Sex

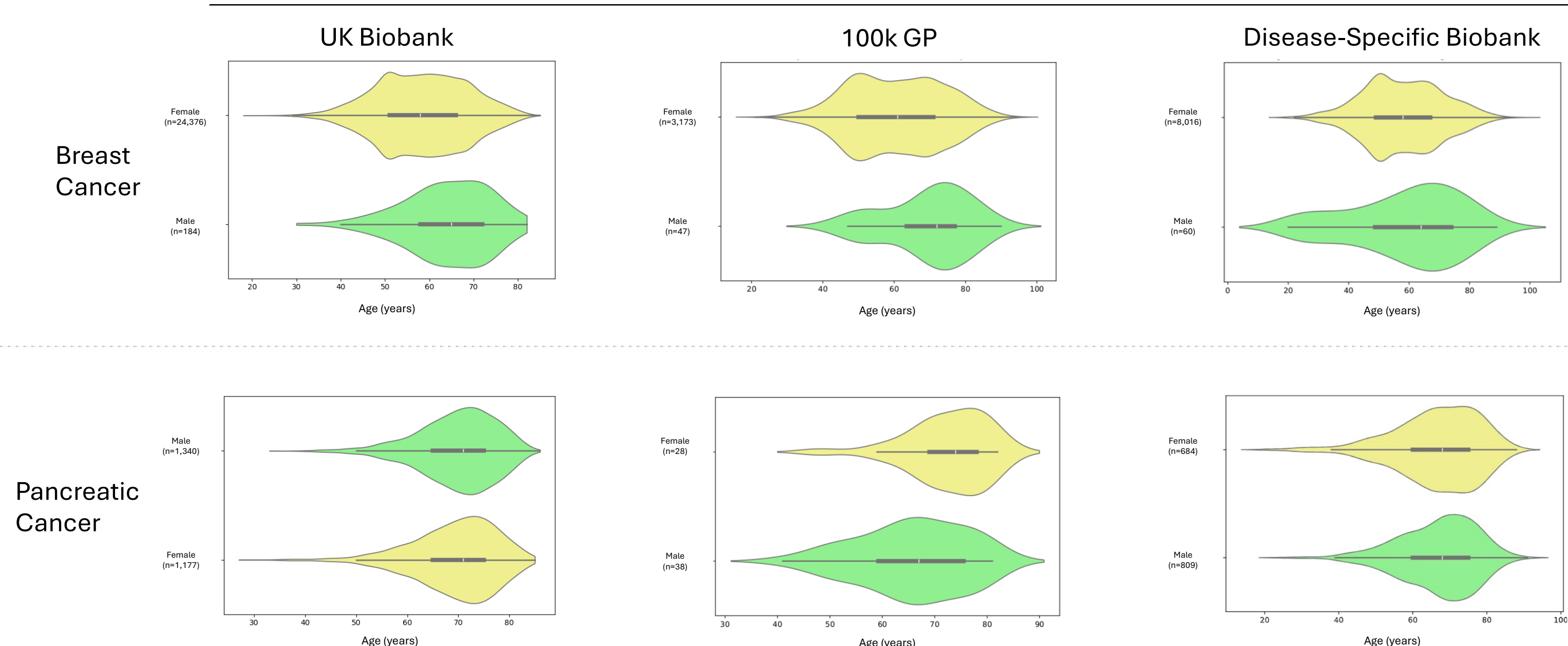


Figure A. Across the breast cancer cohorts, female predominance exceeded 98%. Male breast cancer cases exhibited later median age at diagnosis than females, with higher proportions of hormone receptor-positive disease. There are no significant sex-specific differences in age at diagnosis or age distribution patterns in the pancreatic cohort.

B. Ethnicity

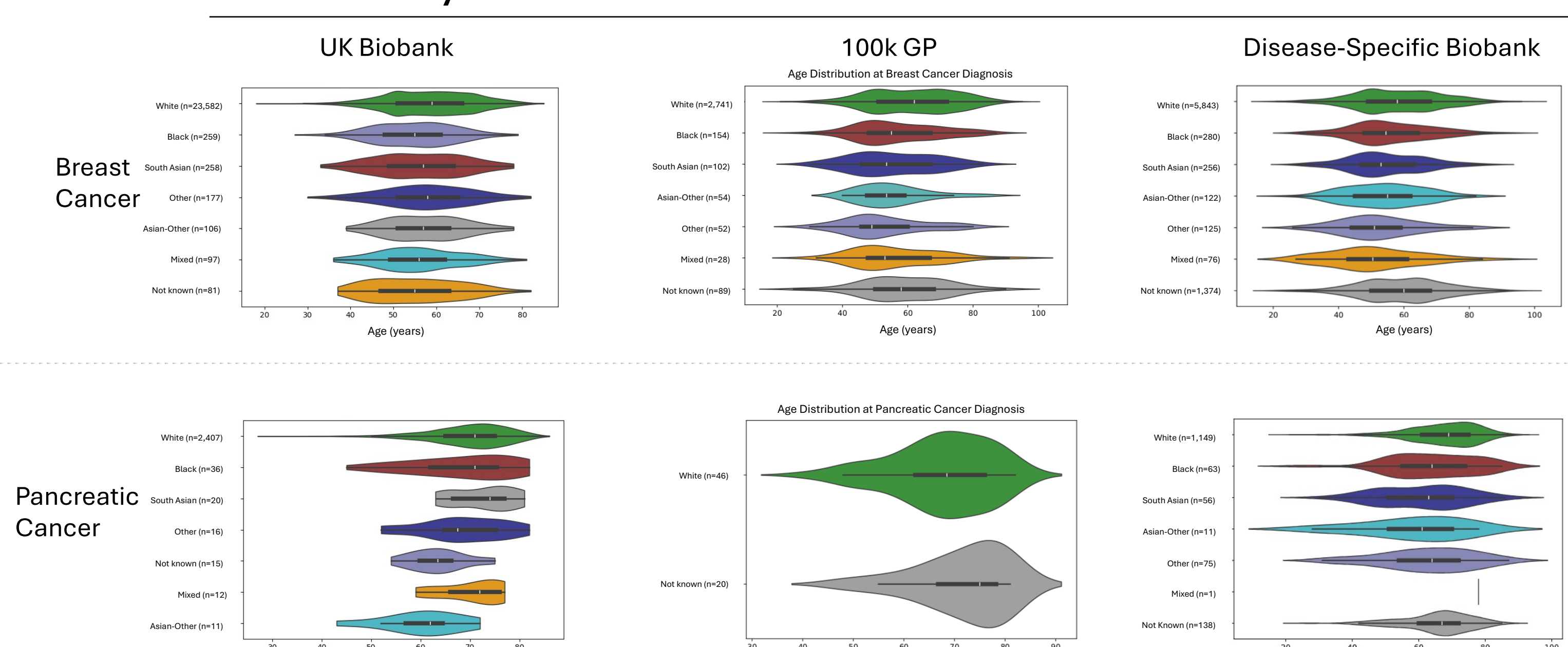


Figure B. Younger median ages at diagnosis were consistently observed among ethnic minority participants relative to White participants, irrespective of cancer.

C. Biospecimens and Data

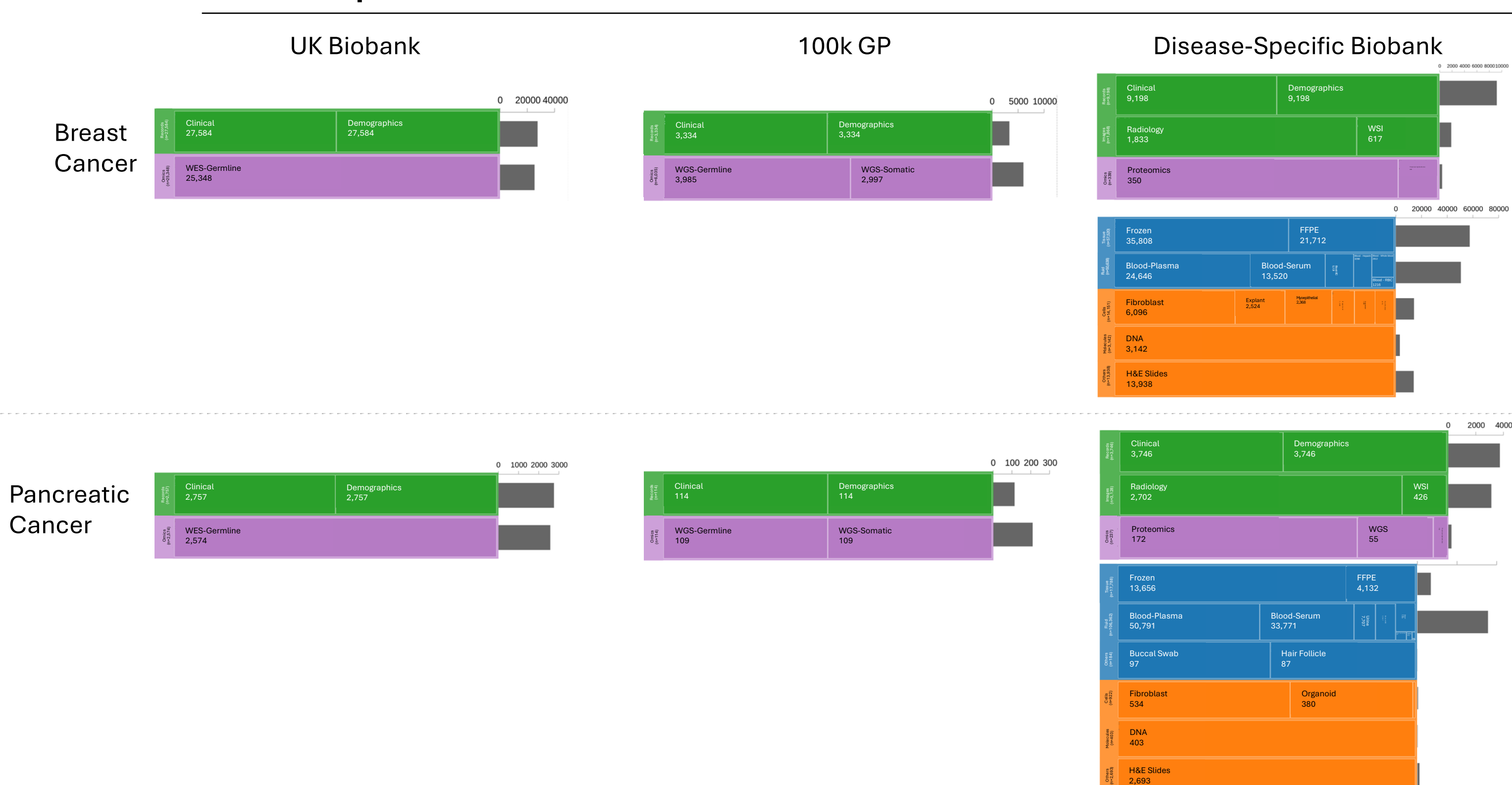


Figure C. These resources provide access to a wealth of complementary biospecimens, longitudinal health records and imaging and molecular data.

