MRC Epidemiology Unit

UNIVERSITY OF CAMBRIDGE

epic
European Prospective Investigation of Cancer

The EPIC-Norfolk study: 15 years of genomic discoveries.

A history of genomic studies and the role of EPIC-Norfolk

Genome-wide studies are a particular type of genetic analysis. We assess variations across the whole length of a person's DNA, and try to identify those that are more common in people with a disease or risk factor. The EPIC–Norfolk study has been involved in many of these.

2007 Wellcome trust case-control consortium publishes papers on 7 common diseases.

2008 Genome-wide EPIC data is used to identify the gene *MC4R*, the second to be associated with BMI.

2010 The 1000 Genomes Project expands the number of variants studied to ~15 million.

2012 EPIC contributes to a study that finds 65 Type 2 diabetes regions, two of which have sex-specific effects.

2016 Sequencing studies, which measure all the base-pairs not just markers, start to report results.

2019 A Genome-wide study of birthweight highlights the links to adult diseases.

2022 The final gaps in the human genome project are identified meaning a fully complete human genome is now available.

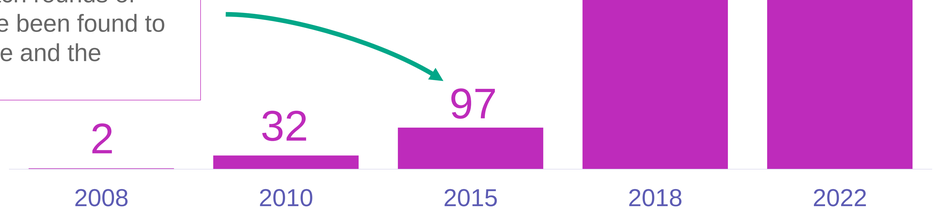
2022 A study of height in 5.4 million people finds genetics explains 40% of the trait.

EPIC-Norfolk contributes to finding thousands of genetic signals and revealing biology

The use of genome-wide data in studies with rich phenotyping, such as EPIC-Norfolk, has led to increasingly large numbers of genetic signals associated with measures of health.

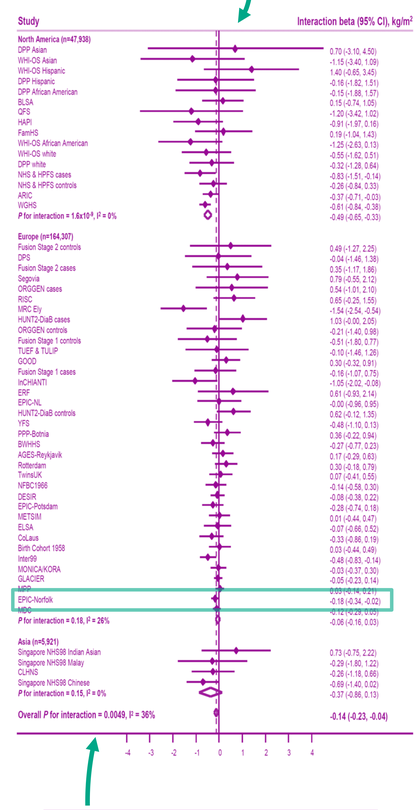
Since 2008 (when the first genome-wide study using EPIC data was published) EPIC has contributed to over 75 such studies.

Increasing numbers of genetic signals linked to BMI have been discovered in each rounds of analysis. Nearly all of these have been found to act in the brain to control appetite and the response to food.



EPIC-Norfolk as part of large-scale international efforts

In order to find as many genetic regions as this, we need to combine a large number of studies of which EPIC is one of the largest.



"Forest" plots show the consistency of an association across different studies. Note that a robust signal might not be apparent in any one alone.

Exploring the causes of Type 2 diabetes

A common use of genetic data is to assess causal associations. The wealth of genetic data on Type 2 diabetes has led to efforts to identify modifiable risk factors.

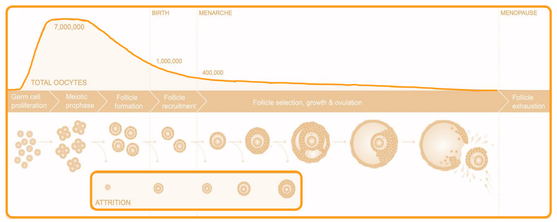
What increases risk:
Depression,
Insomnia,
Faster heart rate,
Obesity, Smoking,
Higher systolic blood pressure.

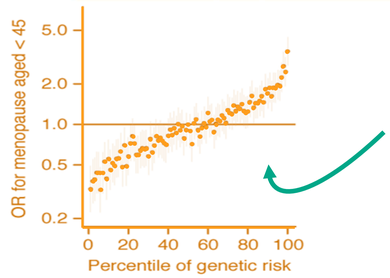
What decreases risk:
Later age at menarche,
Greater birthweight,
Taller height.

We call these studies "Mendelian randomisation", as they take advantage of the random allocation of genes in eggs and sperm to mirror a randomised clinical trial.

Expanding the understanding of menopause

Genetic studies have informed our understanding of the age women go through menopause, with the identified loci linked to protecting the DNA in oocytes (egg cells).





In the future, we might be able to use genetics to identify women who are more likely to have early menopause. This could potentially inform family planning.

Investigating the causes and prevention of diabetes and obesity

www.mrc-epid.cam.ac.uk