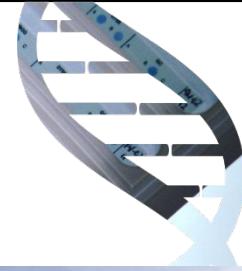


---

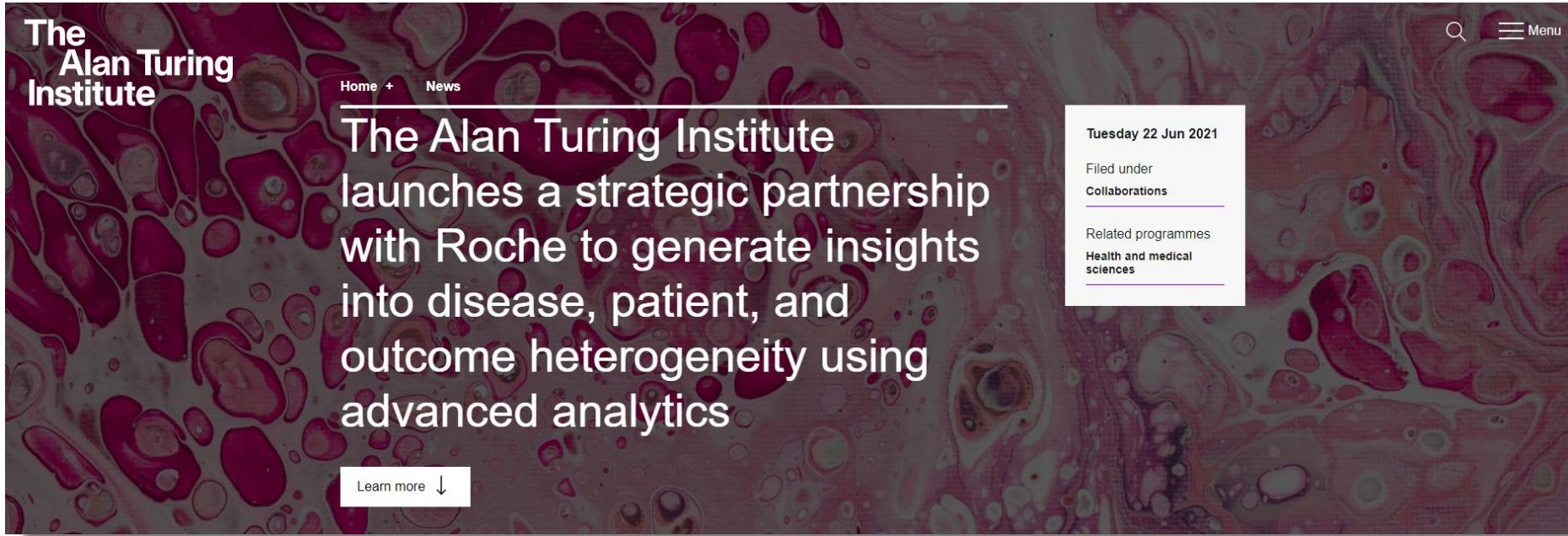
# Turing-Roche Strategic Partnership

## Knowledge Sharing Session

### *January 23rd 2023*



# A Big Announcement in June 2021



The Alan Turing Institute

Home + News

## The Alan Turing Institute launches a strategic partnership with Roche to generate insights into disease, patient, and outcome heterogeneity using advanced analytics

[Learn more ↓](#)

Tuesday 22 Jun 2021

Filed under

Collaborations

---

Related programmes

Health and medical sciences

What's Been Happening Since Then?  
What Happens Next?

# Our Partnership Team



**Professor Ben  
MacArthur**

Director of AI for Science and  
Government, Deputy  
Programme Director for Health  
and Medical Sciences, and  
Turing Fellow



**Professor Chris  
Holmes**

Programme Director for Health  
and Medical Sciences, and  
Turing Fellow



**Katrina Payne**

Partnerships Development Lead



**Ryan Copping**

Global Head of Data Science  
Acceleration, Product  
Development, Roche



**Sarah McGough**

RWD Analytics,  
Roche



**Chris Harbron**

Expert Statistical Scientist and  
Advanced Analytics Lead,  
Roche



**Alvaro Sahun**

Research Project Manager



**Vicky Hellon**

Community Manager, Turing-  
Roche Partnership



**Michelle O'Gorman**

Project Management,  
Roche



**Jay Dougherty**

Alliances,  
Roche

## Our Agenda Today

An Introduction To the Turing Institute and Roche, its Research, and its Data

**Roche-Turing Strategic Partnership:  
Our North Star & Overview**

**Roche-Turing Strategic Partnership:  
Our Post Docs**

**Roche-Turing Strategic Partnership:  
Our Research Themes & Projects**

**Roche-Turing Strategic Partnership:  
Our Community Activities**

# The Alan Turing Institute

## ***UK National Institute for Data Science and AI***

- Network of university, industry, charity, government partners
- Strategic government investment
- Headquarters in the British Library
- Directed by Prof. Sir Adrian Smith (PRS)



# Turing Institute's Strategic and Project Partners

The  
Alan Turing  
Institute 



BILL &  
MELINDA  
GATES  
foundation



Lloyd's Register  
Foundation

GCHQ  
[dstl]

Ministry  
of Defence

NATS

Office for  
National Statistics



AUTISTICA

BRITISH  
AIRWAYS

British  
Antarctic Survey  
NATIONAL ENVIRONMENT RESEARCH COUNCIL

British Heart  
Foundation

Cystic  
Fibrosis Trust

Department for  
Digital, Culture,  
Media & Sport



FCA  
FINANCIAL CONDUCT AUTHORITY

GREAT  
ORMOND  
STREET  
HOSPITAL  
CHARITY

GREATER  
LONDON  
AUTHORITY

HDRUK  
Health Data Research UK

The  
Health  
Foundation

IMPROBABLE

ICO.



LLOYDS  
BANKING  
GROUP

Met Office

Microsoft

NHS  
SCOTLAND

Alzheimer's  
Research  
UK

Office for  
Artificial  
Intelligence

Ofcom

Public Health  
Scotland

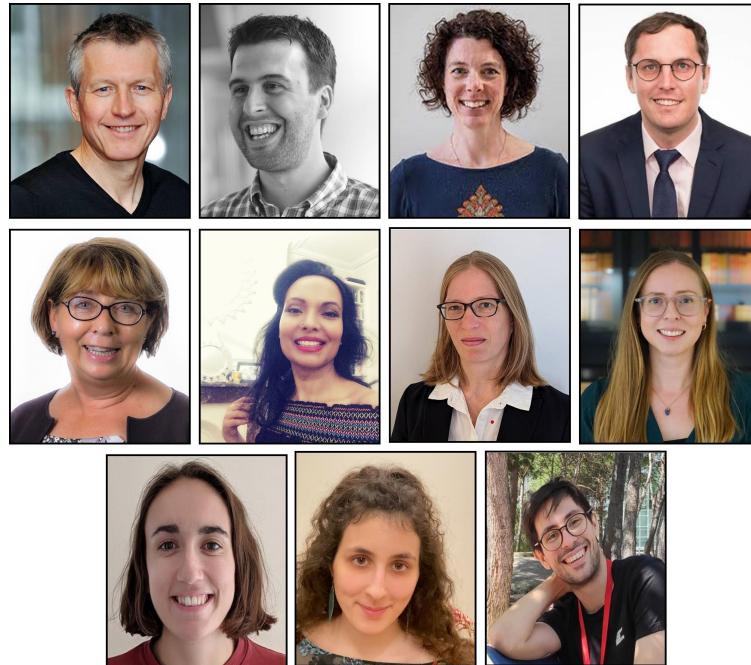
ROLLS  
ROYCE

TOYOTA  
mobility  
FOUNDATION

UN  
DP

# Turing Institute Health Programme

- Started in 2018
- Chris Holmes, director
- Ben MacArthur, deputy programme director (2020)
- Alisha Davies, ASG health lead
- Owen Rackham, lead for cell and molecular medicine
- Programme, project and community managers

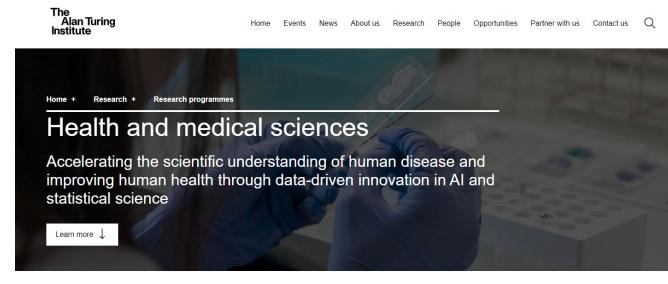


# Health & Medical Sciences Programme: accelerating the understanding of human disease & improving human health

Supporting the rigorous, robust, development and application of AI and data science to problems in human health

Building long term **strategic relationships** with government, charity, & industry partners

Tackling major real-world applications with domain experts, which in turn informs methodological and theoretical innovation



## Introduction

AI and data science will improve the detection, diagnosis, and treatment of illness. They will c health service providers to anticipate demand and deliver improved patient care.

The Turing's health programme will deliver a schedule of major research into the theory and r underpinning medical and health applications that will enable scientists to do better science, i patient trust.



# Health Programme Current Areas of Focus



- Tackling major application areas including **early detection of disease, risk prediction, pandemic response and future preparedness, multimorbidity, and molecular medicine**
- Working in close partnership with Turing programmes in Defence and Security, Public Policy, Tools Systems and Practices, and AI for Science
- Core agenda in robustness, accurate uncertainty quantification, and reproducible research for AI for Health

# The Roche Group

*A leading healthcare company dedicated to sustainable innovation*

**1896**

Founded in Basel  
Family still holds  
majority stake



*Lasting...*

>30 medicines  
on World Health  
Organization List of  
Essential Medicines



**#1**

**R&D investor** in  
healthcare



**58.3 bn**

Sales 2020  
(CHF)



**101'465**

Employees worldwide

**28,900,000**

people treated  
worldwide with our  
medicines in 2020



*... and Sustainable*

Among top companies in **Dow Jones  
Sustainability Index (Life Sciences  
Sector)** for **11** consecutive years

**37**

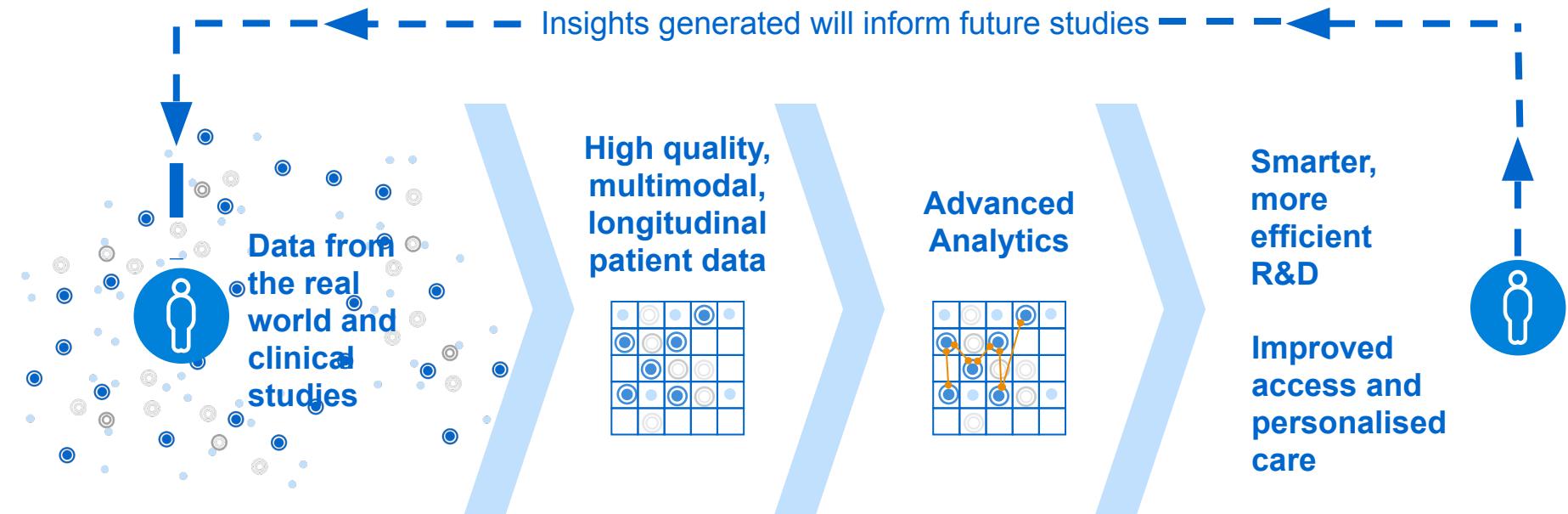
FDA Breakthrough  
Designations

<sup>1</sup> Genentech became a full member of the Roche group in March 2009

*Personalised healthcare means better health, at a lower cost, for people and society. It means shifting from a one-size-fits-all approach to the best care for each person.*

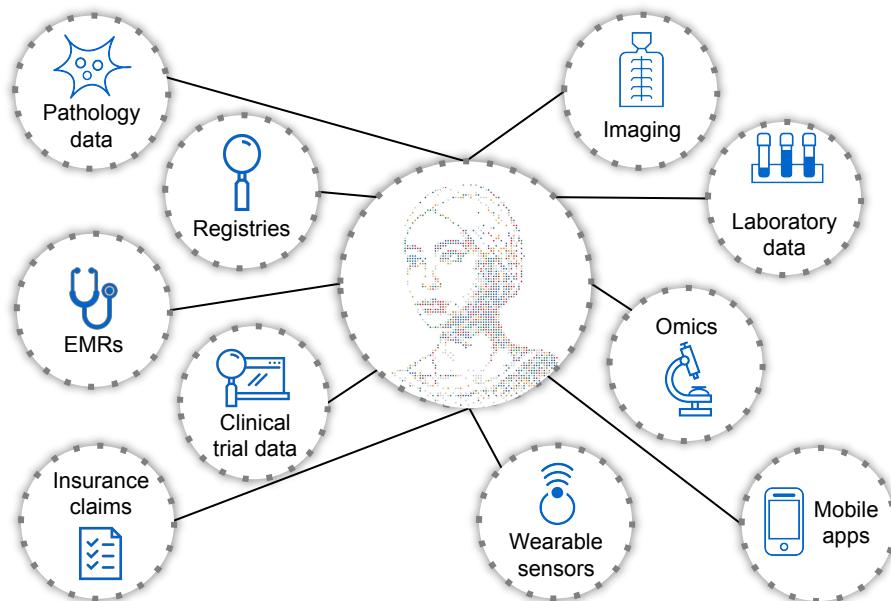


# Data & Advanced Analytics (AA) are key enablers to transform healthcare



***Focus:** Applying advanced analytics will enable understanding of patient and disease heterogeneity and its relevance to clinical outcomes at an unprecedented resolution*

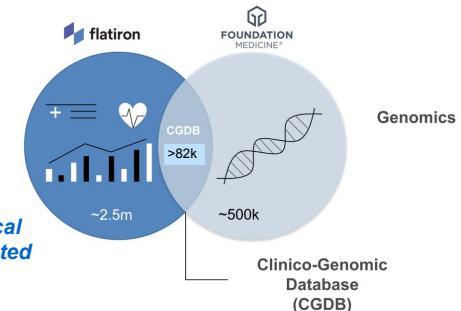
# Roche is integrating high quality, longitudinal, multimodal healthcare data



## Examples:

1) Multiple leading research grade Real World Data (RWD) assets, including CGDB

Electronic Health Records & Outcomes



2) Data from >220k patients from our clinical trials (& some observational cohorts) curated into disease area data marts

Enhanced Data and Insights Sharing (EDIS)

Transforming culture & technology infrastructure to generate disease-specific datamarts and share across physical and organisation limits

Making data F.A.I.R  
Findable. Accessible. Interoperable. Reusable.<sup>1</sup>

Broader and purposeful availability of aggregated data to accelerate insights generation

Cancer immunotherapy      Autism spectrum disorder      Non-Hodgkin's lymphoma

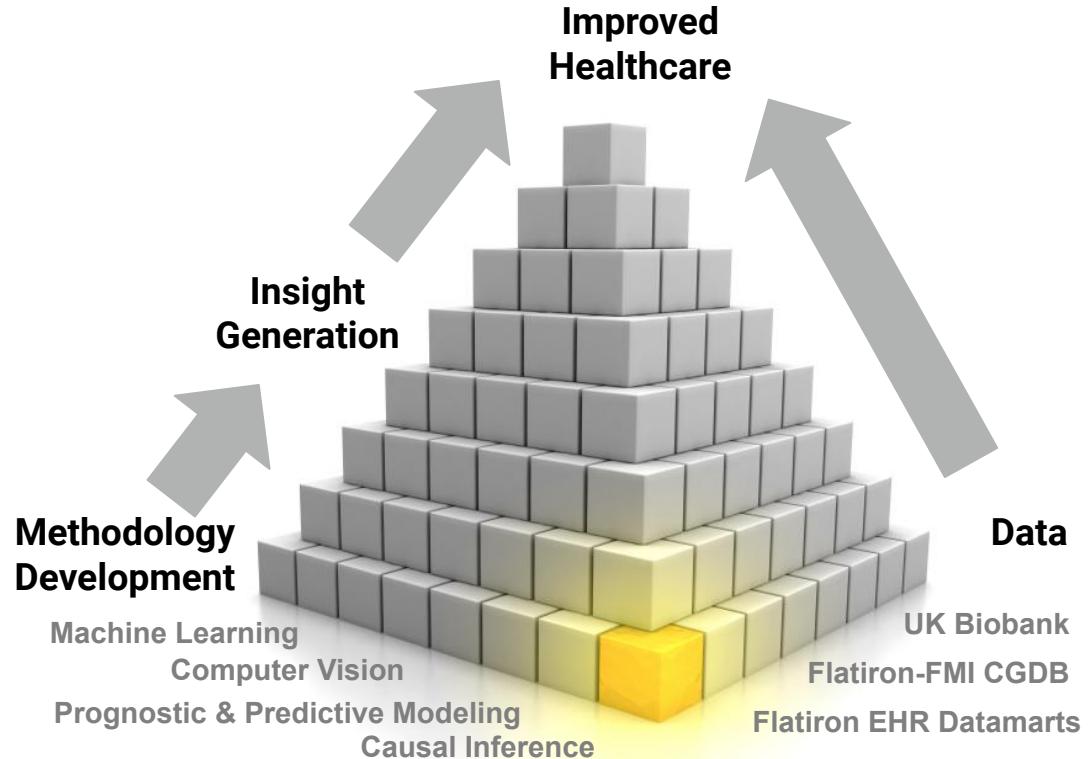
Asthma      Alzheimer's Disease      Triple Negative Breast Cancer

*This will enable a deeper understanding of each patient and their disease*

# Turing-Roche Strategic Partnership

## *Tackling Our North Star in an iterative, stepwise way*

***To enable the generation of insights to better understand patient and disease heterogeneity and its relevance to clinical outcomes at an unprecedented level of precision in order to improve clinical care***



# Turing-Roche Partnership Principles & Opportunities

## Open Science



Jointly building exciting science together

## Data in Volume, Depth & Quality



Access to wide diversity of data sources from both parties

## Collaborative



Active contributions from both Roche & Turing

## Working across communities



Wide and diverse set of talent and experiences across the two partners

## Prioritisation



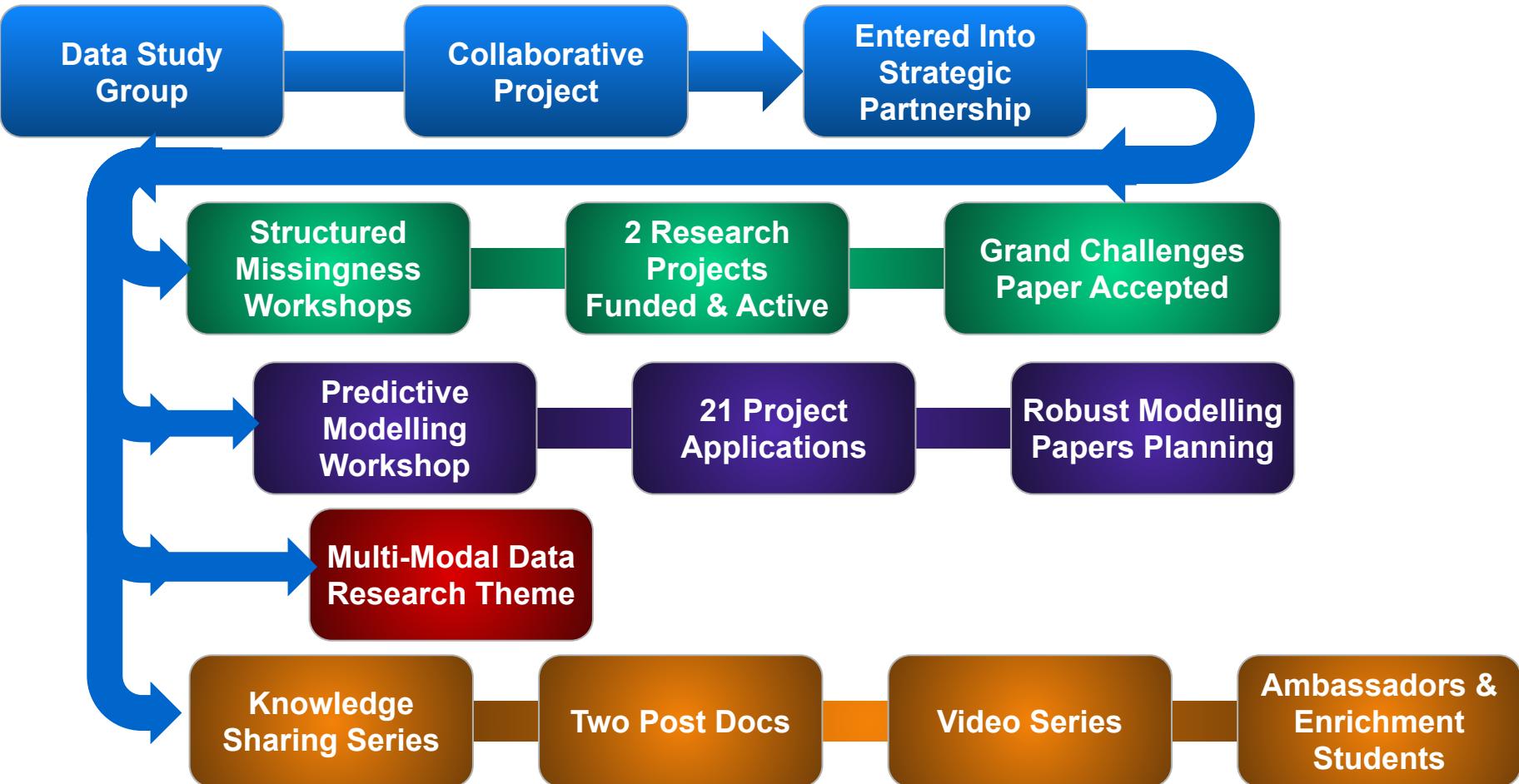
Agile operating model and abundance of questions

## North Star



Will drive further fundamental research and have an impact on patients

# Turing-Roche Partnership Journey



# Introducing Our Senior Research Associates

Tapabrata (Rohan) Chakraborty

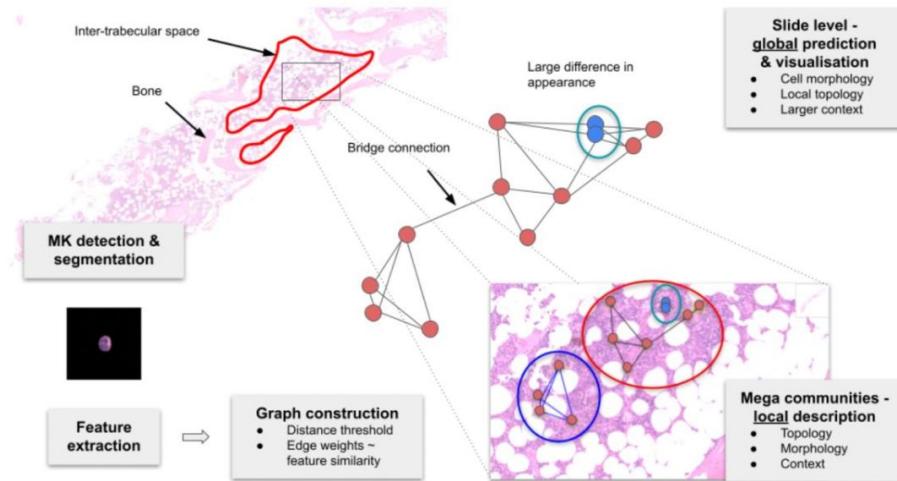


Chris Banerji

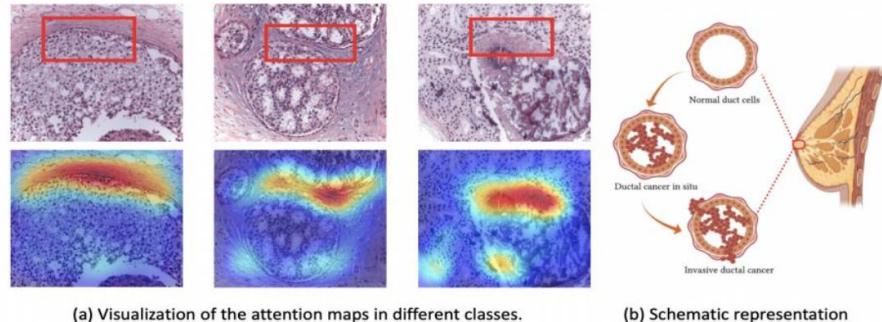


# Rohan's Intro: SRA at Turing-Roche, previously postdoc at Oxford, earlier PhD in CS, focus on explainable AI/ML models for imaging.

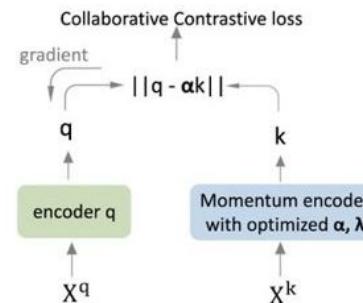
## Graph Network based interpretable models



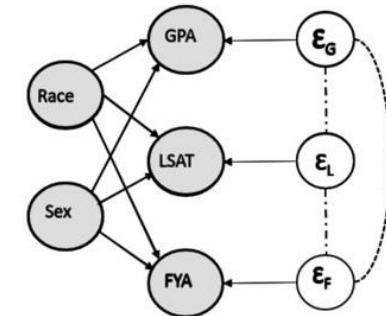
## Attention based interpretable models



## Continual Learning

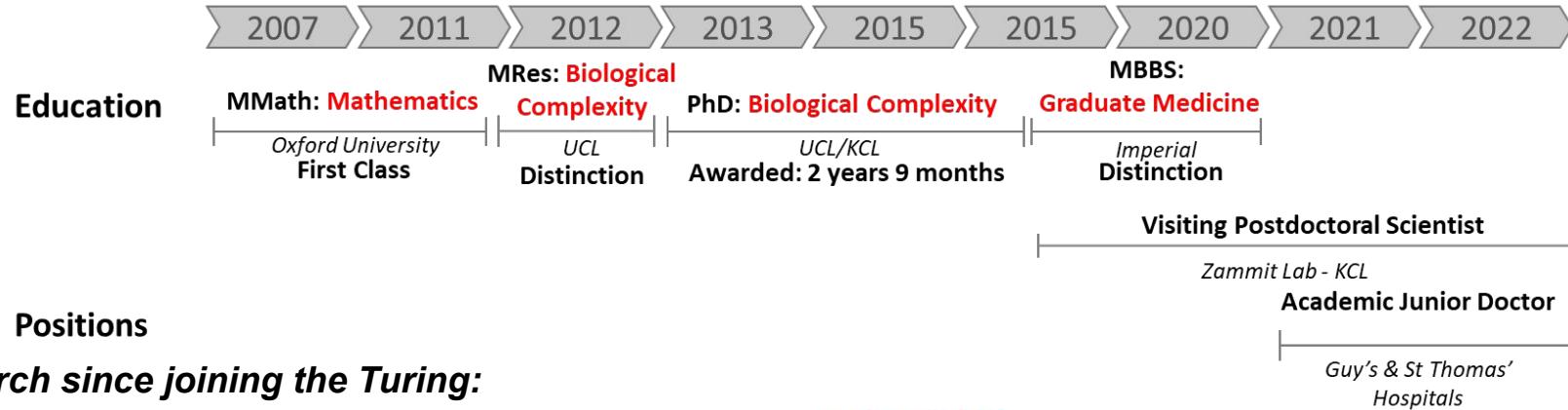


## Causal Inference



# Chris Banerji

## Road to Turing-Roche:



## Positions

### Research since joining the Turing:

#### Transcriptomic Biomarkers

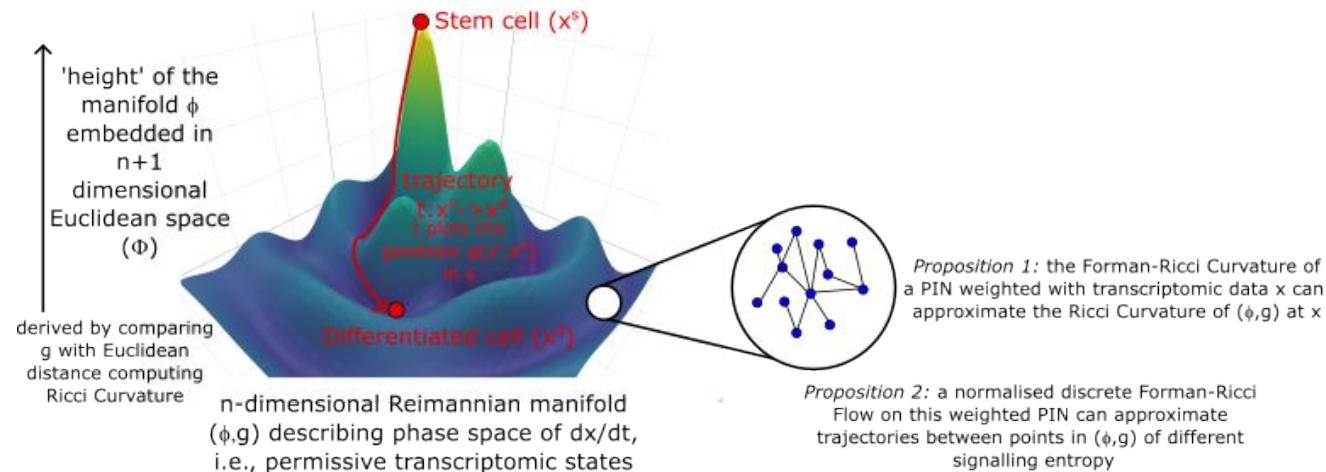
<https://www.biorxiv.org/content/10.1101/2022.08.31.506017v1.abstract>

#### Stochastic Gene Expression

<https://www.biorxiv.org/content/10.1101/2022.12.12.520053v1.abstract>

#### 'Rules' of cellular differentiation and oncogenesis

*Network theory and differential geometry*



# We Have Prioritised Three Major Research Themes

## Structured Missingness

### Projects

- Developing a coherent Bayesian modelling and imputation framework that accounts for, and utilises, Structured Missingness
- Using network science to quantify the geometry of “missingness”

#### Learning from data with structured missingness

Robin Mitra<sup>1,2,\*</sup>, Sarah F. McGough<sup>3,\*</sup>, Tapabrata Chakraborti<sup>1</sup>, Chris Holmes<sup>1,4</sup>, Ryan Copping<sup>2</sup>, Niels Hagenbuch<sup>5</sup>, Stefanie Biedermann<sup>6</sup>, Jack Noonan<sup>7</sup>, Brieuc Lehmann<sup>2</sup>, Aditi Shenvi<sup>8</sup>, Xuan Vinh Doan<sup>9</sup>, David Leslie<sup>1</sup>, Ginestra Bianconi<sup>1,10</sup>, Ruben Sanchez-Garcia<sup>1,11</sup>, Alisha Davies<sup>1,12,13</sup>, Maxine Macintosh<sup>1,14</sup>, Eleni-Rosalina Andrinopoulou<sup>15</sup>, Anahid Basiri<sup>1,16</sup>, Chris Harbron<sup>17,\*</sup>, and Ben D. MacArthur<sup>1,11,18,\*</sup>

## Predictive Modelling

### Themes

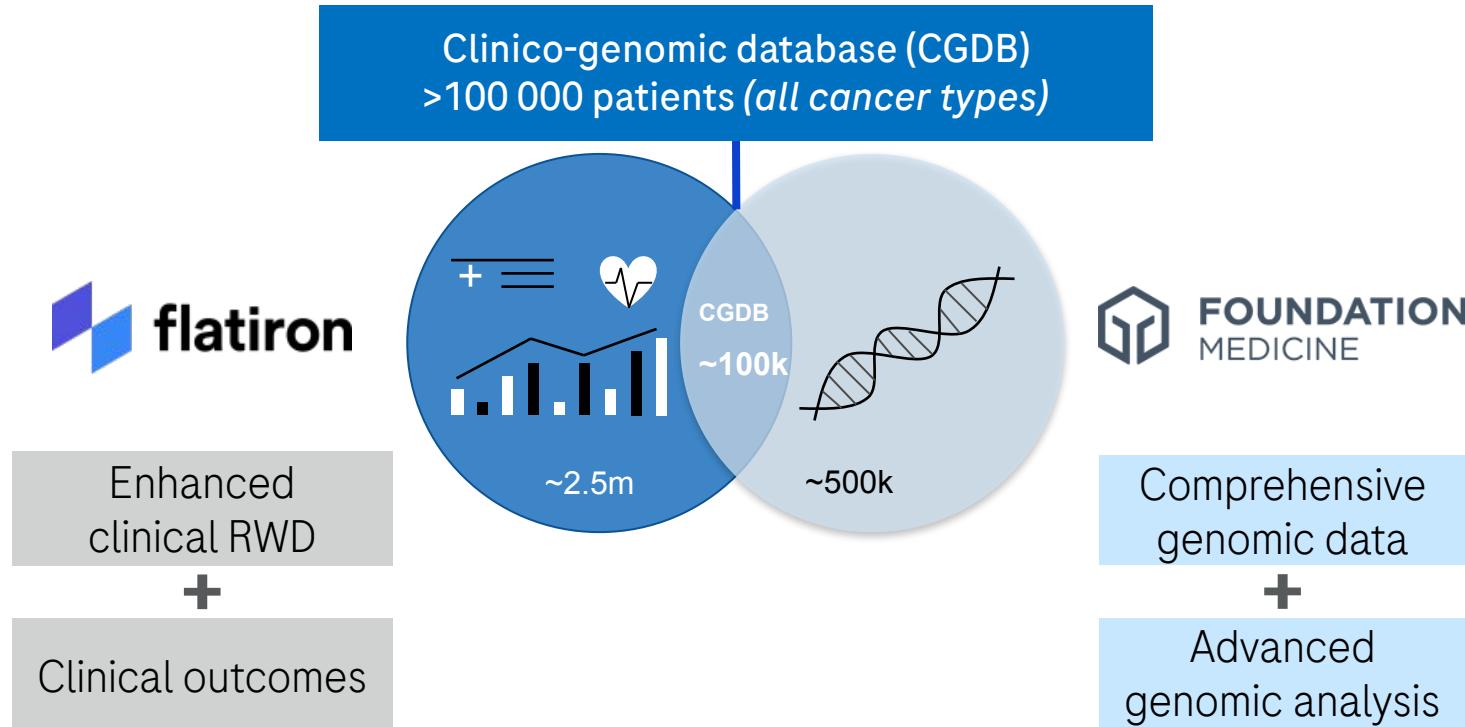
- Generalisation despite data heterogeneity
- Prediction uncertainty in personalised healthcare
- Algorithmic explainability and fairness for AI in healthcare

## Multimodality

- Prioritising combining Genomics + Imaging

## Theme 1: Structured Missingness example - CGDB

The Clinico-Genomic Database links Flatiron electronic health records with Foundation Medicine (FMI) comprehensive genomic profiling for tens of thousands of cancer patients in the U.S.



# Structured Missingness Projects

Developing a coherent Bayesian modelling and imputation framework that accounts for and utilises Structured Missingness



Dr Robin Mitra



Professor Ana  
Basiri



Dr Eleni-Rosalina  
Andrinopoulou

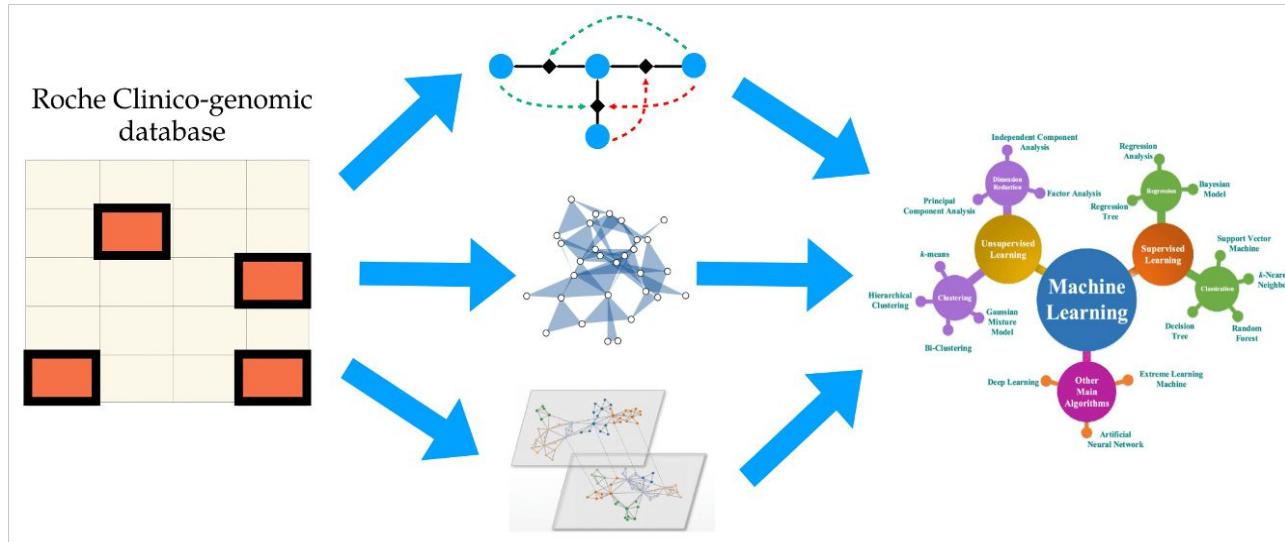


Dr James Jackson

- In large complex data sets combined from multiple sources, Structured Missingness (SM) is likely. I.e. missing values arise within an inherent association.
- Imputation is an appealing approach but may not be appropriate for some types of SM, e.g. for missing PSA scores in female patients.
- Our project will develop a Bayesian modelling framework that can deal with SM, imputing missing values only when this is possible to do.
- We will also investigate suitability of existing packages for missing data, e.g. MICE in R, to deal with SM, and consider modifications to these where appropriate.
- We also seek to use missing values as information itself to help better inform the imputations.

# Structured Missingness Projects

Using network science to quantify the geometry of “missingness”



①

Characterize pairwise and higher-order interactions

②

Extract Structural Missingness Networks (SMNs)

③

Characterize the Geometry of SMNs



Professor Ginestra Bianconi



Dr Ruben Sanchez Garcia

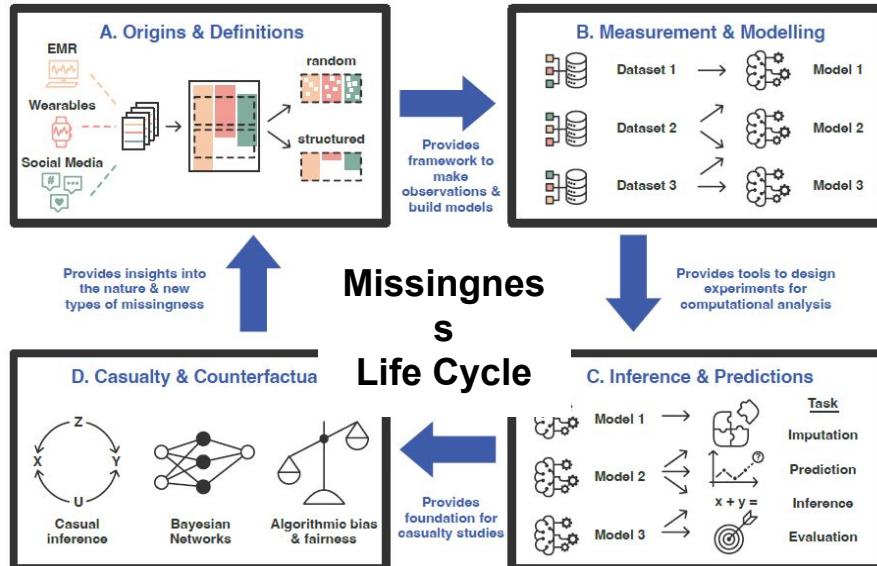


Dr Anthony Baptista

# Structured Missingness Publication

## Learning from data with structured missingness

Robin Mitra<sup>1,2,\*</sup>, Sarah F. McGough<sup>3,\*</sup>, Tapabrata Chakraborti<sup>1</sup>, Chris Holmes<sup>1,4</sup>, Ryan Copping<sup>5</sup>, Niels Hagenbuch<sup>5</sup>, Stefanie Biedermann<sup>6</sup>, Jack Noonan<sup>7</sup>, Brieuc Lehmann<sup>2</sup>, Aditi Shenvi<sup>8</sup>, Xuan Vinh Doan<sup>9</sup>, David Leslie<sup>1</sup>, Ginestra Bianconi<sup>1,10</sup>, Ruben Sanchez-Garcia<sup>1,11</sup>, Alisha Davies<sup>1,12,13</sup>, Maxine Macintosh<sup>1,14</sup>, Eleni-Rosalina Andrinopoulou<sup>15</sup>, Anahid Basiri<sup>1,16</sup>, Chris Harbron<sup>17,\*</sup>, and Ben D. MacArthur<sup>1,11,18,\*</sup>



- Paper arose from workshops and subsequent discussions, and is written by a multidisciplinary community
- Outlines current approaches to SM and proposes a series of grand challenges that collectively provide a roadmap for the development of SM as a field of study that will help advance machine learning at scale.

# Structured Missingness Publication

## Grand Challenges For Structured Missingness



Defining and Characterising  
Structured Missingness



Exploring SM  
Geometry and Visualisation



Prediction



Inference and Estimation



Causality



The Role of Imputation



Design Considerations



Benchmarking And Evaluation



Ethical Implications

# Theme 2: Predictive Modelling

## Identified 3 Key SubThemes

- Generalisation despite data heterogeneity
- Prediction uncertainty in personalised healthcare
- Algorithmic explainability and fairness for AI in healthcare

## Workshop

- In-person at the Turing
- Attended by ~40 attendees
- Explore subthemes and identify connections and synergies within and between subthemes

## Projects

- Call for funded projects
- Outcomes to be shared once contractual arrangements completed

## Publications Planned

1. Robustness in healthcare predictive models
2. Patient-Centric Fairness

## Theme 3: Multi-Modal Data Integration and analysis

- Initial focus on combining **genomics** and **imaging** data for oncology
- Discussions on best approach to address this are underway
- Theme activities will start in 2023



# Partnership Community Building to date

- **Need:** build community from scratch
- **Approach:** putting the basics in place!
  - Low barrier space for connecting: Slack Workspace
  - Leveraging existing networks to raise awareness
  - Quick wins
- **Development of community strategy and activities**



# Current Community Activities/Resources

## Opportunities To Get Involved

### Slack Channel

#### Turing-Roche Partnership

- # general
- # interesting-reading-and-resources
- # introductions
- # jobs-and-opportunities

### Monthly Newsletter



#### The Alan Turing Institute – Roche strategic partnership

Welcome to December's Turing-Roche newsletter! Wishing you a restful upcoming holiday season. Here you can find what's new in the Turing-Roche strategic partnership and opportunities, events and news from both The Alan Turing Institute and Roche.

### GitHub Repository

[View repository](#)

The Turing-Roche Strategic Partnership

Welcome to community repository for the Turing-Roche Strategic Partnership. Here you'll find project and community related documentation from our partnership.

### Knowledge Share Series

The Turing-Roche knowledge share series

An event series for Turing-Roche partnership updates, knowledge sharing and new perspectives

### Partnership Advisors



Professor Alejandro Frangi  
Turing Fellow



Dr Michael Barnes  
Turing Fellow



Dr Florian Markowetz  
Senior Group Leader, Cancer Research UK  
Computational Neuroscience at the University of Cambridge and Turing University Ltd - Cambridge



Professor Zoi Kourouzi  
Turing Fellow, Principal Investigator of Cognitive Computational Neuroscience at the University of Cambridge and Turing University Ltd - Cambridge



Professor Heather Harrington  
Professor of Mathematics, University of Oxford



Dr Dan Ruderman  
Director of Digital Pathology for Oncology Biomarker Development, Genentech



Dr Astrid Kiernan  
Leader for Personalized Healthcare in early research and development, Roche



Dr Marius Garmhausen  
Principal Data Scientist, Personalised Healthcare Analytics, Roche

### YouTube Channel

#### The Turing-Roche Strategic Partnership

@turing-roche  
36 subscribers

HOME

VIDEOS

PLAYLISTS

COMMUNITY

CHANNELS

Problem Statement

This video is a call for a young researcher at RocheGeno to solve a problem using their skills and expertise. The researcher is asked to identify a specific challenge in the field of genomics and propose a solution. The video ends with a question: "What is the most likely origin of this patient?"

Demystifying Bayesian Estimation with Dr Tapabrat...

Academic and Industry Perspectives on Data Scien...

# Turing-Roche Community in 2023

## Opportunities To Get Involved

- Early Career Researcher Ambassador Scheme
- Planning new events- partnership mini symposium
- Facilitating connections outside of ‘formal’ partnership e.g Turing Way
- And any other ideas from the community, would love to hear from you!



**For more information or to sign up for the newsletter,  
emails and the slack channel, please visit :**

<https://www.turing.ac.uk/research/research-projects/alan-turing-institute-roche-strategic-partnership>



**Any Questions**