Prof. Paul O'Reilly

Labsite: https://labs.icahn.mssm.edu/oreillylab/

EMPLOYMENT

2024 – Professor (Tenured) in Statistical Genetics, Icahn School of Medicine, Genetics and Genomic Sciences, Mount Sinai, NYC

2019 – 24 Associate Professor in Statistical Genetics, Icahn School of Medicine, Genetics and Genomic Sciences, Mount Sinai, NYC

2017 – 19 Reader (Associate Prof) in Statistical Genetics, King's College London, IoP

2013 – 17 Senior Lecturer in Statistical Genetics, King's College London, IoP

2011 – 13 Lecturer (Assistant Prof) in Genetic Epidemiology, Imperial College London, Dept. Epidemiology and Biostatistics

2012 Honorary Assistant Professor, University of Hong Kong (HKU).
3-month sabbatical to Prof Pak Sham's group, Dept of Psychiatry (HKU)

2008 – 2011 Postdoctoral Research Associate, Imperial College London

EDUCATION

2009 PhD Population Genetics, Imperial College London
Detecting genetic loci under recent positive selection
1st Supervisor: Prof David Balding; 2nd supervisor, Dr Ewan Birney

2004 MSc Genetic Epidemiology, The University of Sheffield, UK
Supervisor: Prof Chris Cannings

2003 BSc Mathematics, The University of Leeds, UK
Grade: 1st (equivalent to Summa Cum Laude)

GRANTS

2024 R01 NIA **(PI)** ($3.4M; 20% salary): *Pathway-based polygenic risk scores for sub-typing Alzheimer's disease.* (13% score; pending).

2023 R01 NHGRI **(PI)** ($3.2M; 40% salary): *BridgePRS: bridging the gap in polygenic scores between ancestries.*

2020 R01 NIMH **(PI)** ($3.1M; 35% salary): *Next-generation, pathway-specific, polygenic risk scores*

2020 U01 NHGRI (Co-I) ($3.6M; 10% salary): *Genomic risk in clinical care to promote health equity in New York City patients*

2020 R01 NIMH (Co-I) ($2.7M; 5% salary): *Maternal health in pregnancy and autism risk – genetic and non-genetic mechanisms*

2019 Klarman Foundation **(Sub-award PI)** ($750k; 10% salary): *Translation of Anorexia Nervosa variants into genes, pathways and tissues*

2017 EU-ERC Starter (Co-I) (~€1.5M; 10% salary)

2016 MRC Methods Grant **(PI)** (~£500k; 35% salary): *Polygenic score methodology in the emerging field of Polygenic Epidemiology*

2016 Industry-funded (RGA) collaboration grant **(MPI)** (~£120k; 5% salary)
2015 Wellcome Trust Seed Award **(PI)** (~£100k; 10% salary)

RESEARCH PROFILE

Background: Having completed undergraduate and masters training in Mathematics and Genetic Epidemiology, respectively, my PhD focused on developing methods to detect recent positive selection, resulting in a 1st author paper in *Genome Research* that highlighted the intimate link between recombination and selection. I later co-authored a correspondence in *Nature Genetics* on the topic (O'Reilly & Balding, 2011). Subsequently, I shifted focus to Genetic Epidemiology, contributing to numerous high-impact GWAS publications that identified loci across a range of phenotypes. My methodology research since has focused on two areas: multi-trait GWAS and polygenic prediction. In 2012, I published a method, *MultiPhen,* for performing multi-trait GWAS that is one of the leading such methods (**> 400 citations**), which along with a comparison study published in 2017 (Porter & O'Reilly; **> 100 citations**), highlights the benefits of multi-trait GWAS methods. In 2015, I published a method/software for performing polygenic risk score analyses, *PRSice*, which has become the most applied PRS software in the field (**> 2800 citations**), the following year I received a MRC methods grant *Polygenic score methodology in the emerging field of Polygenic Epidemiology*, and in 2020 I was awarded an R01 methods grant *Pathway-specific, function-informed, polygenic risk scores*, followed in 2023 by an R01 methods grant *BridgePRS: bridging the gap in polygenic scores between ancestries.*

Lab summary: My lab develops statistical and computational approaches to understand how human genetic variation, in combination with the environment, leads to disease. Given their proxy for genetic liability itself, our primary focus is on the theory and application of Polygenic Risk Scores (PRS). In 2015, we published the popular PRS software *PRSice*, followed in 2019 by *PRSice-2* (**> 2800 citations**). We also wrote a *Guide to Performing Polygenic Risk Score analyses* (Choi, Mak, O'Reilly. 2020; *Nature Protocols*) that has become one of the most read papers in the field (**> 1800 citations**). We have now extended our PRS software suite to include, *PRSet*, which performs PRS analyses across genomic pathways, and a method, *BridgePRS*, to compute PRS accurately in diverse ancestry populations.

Much (but not all) of the research in my lab falls into one of **4 key themes**:

(1) Pathway-specific, function-informed, polygenic risk scores

(2) Polygenic risk scores for diverse and admixed populations

(3) Investigating genetic architecture in the tails of complex trait distributions

(4) The Statistical Genetics of Complex Traits and Diseases

SELECTED PUBLICATIONS

I have published over 100 articles since my first in 2008 (H-index 71; i10-index 122), cited >30k times. Google Scholar page here.

- Hoggart C, Choi S, Garcia-Gonzalez J, Souaiaia T, Preuss M, **O'Reilly PF**. 2024. BridgePRS leverages shared genetic effects across ancestries to increase polygenic risk score portability. *Nature Genetics* 56: 180-86.
- Choi S, Garcia-Gonzalez J, Ruan Y, Wu HM, Johnson J, Hoggart C, **O'Reilly PF**. 2023. PRSet: Pathway-based polygenic risk score analyses and software. *Plos Genetics.* doi.org/10.1371/journal.pgen.1010624
- Choi S, Mak TSH, Hoggart C, **O'Reilly PF**. 2023. EraSOR: Erase Sample Overlap in polygenic score analyses. *Gigascience 12, giad043.*
- Iyegbe C & **O'Reilly PF**. 2022. Genetic origins of schizophrenia find common ground. *Nature* 604: 433-435.
- Socrates A, Maxwell J, Glanville K,…, **O'Reilly PF**. 2021. Investigating the effects of genetic risk of schizophrenia on behavioural traits*. NPJ Schizophrenia* 7: 1-9.
- Wu HM, Goate AM, **O'Reilly PF**. 2021. Heterogenous effects of genetic risk for Alzheimer's disease on the phenome. *Translational Psychiatry.* 11: 1-9.
- Choi SW, Mak TSH, **O'Reilly PF**. 2020. A guide to performing polygenic risk score analyses. *Nature Protocols* 15: 2759-72.
- Choi, SW, **O'Reilly PF**. 2019. PRSice-2: Polygenic Risk Score software for biobank-scale data. *Gigascience*. 8(7) giz082.
- Selzam S, …, **O'Reilly PF**, Plomin R. 2019. Comparing within- and between-family polygenic score prediction. *Am.J.Hum.Gen* 105: 351-63.
- Pingualt JB, **O'Reilly PF** et al. 2018. Using genetic data to strengthen causal inference in observational research. *Nature Reviews Genetics*
- Krapohl E, ..., **O'Reilly PF**, Plomin R. 2018. Multi-polygenic score approach to trait prediction. *Molecular Psychiatry*. **23**: 1368-74.
- Porter HF and **O'Reilly PF**. 2017. Multivariate simulation framework reveals performance of multi-trait GWAS.. *Scientific Reports*
- Euesden J, Lewis CM, **O'Reilly PF**. 2015. PRSice: Polygenic Risk Score software. *Bioinformatics*. 31:1466-8.
- al Basatena NKS, Hoggart, CJ, Coin LJ, **O'Reilly PF**. 2013. The effect of genomic inversions on estimation of population genetic parameters. *Genetics* 193: 243-53.
- **O'Reilly PF** et al. 2012. *MultiPhen*: joint model of multiple phenotypes increases discovery in GWAS. *PLoS ONE.*
- **O'Reilly PF** & Balding DJ. 2011. Admixture mapping provides insights into recombination. *Nature Genetics* 43: 819.
- Wain*, Verwoert*, **O'Reilly PF*** et al. 2011. Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. *Nature Genetics*. 43: 1005-11.
- **O'Reilly PF**, Coin LJ, Hoggart CJ. 2010. invertFREGENE: software for simulating inversions in population genetic data. *Bioinformatics*. 26: 838-40.

- **O'Reilly PF**, Birney E, Balding DJ. 2008. Confounding between recombination and selection and the Ped/Pop... *Genome Research* 18: 1304-13.

TEACHING

I have a huge amount of teaching experience, both in terms of contributing teaching across a multitude of courses at all levels (1:1 high school mathematics personal tutoring, through statistics teaching of trainee clinicians, to statistical genetics modules in MSc/PhD programs), as well as designing and directing of dedicated workshops, modules and MSc programs. Selected examples below:

- PAPG Module Director, Mount Sinai Graduate School (2022-)
- PRS Workshop Co-Director, University of Tokyo, Japan (2024)
- PRS Workshop Co-Director, Makerere University, Uganda (2023)
- PRS Workshop Director (organised by H3Africa), University of Mauritius (2022)
- Co-Designer CEYE Bioinformatics Program (2020)
- PRS Summer School Director, King's College London (2018, 2019)
- MSc Remote Learning Module Director, King's College London (2017-19)
- MSc Deputy Director (*Genes, Environment..*), King's College London (2014-17)
- MSc Module Director (*Genes, Environment..*), King's College London (2014-17)

OTHER SELECTED PROFESSIONAL ACTIVITIES

- Founder of Mount Sinai Data Commons: Data Ark (https://tinyurl.com/yjc573ep)
- Founder of Open Science start-up: *Animatus* (https://animatus.science)
- Co-Founder of Mount Sinai *STOP COVID NYC* phone app and study (details here)
- Co-Designer of Mount Sinai CEYE Bioinformatics Program
- Organiser of the South of England Genetic Epidemiology Group (SEGEG), comprising over 600 UK researchers in the field (2015-19)
- Chair *Faculty Mentoring Committee* of Dept. (GGS)
- Member of *Faculty Recruitment Committee* of Dept. (GGS)
- Member *Collaborations and Initiatives Committtee* of Dept. (GGS)
- Special Statistical Reviewer for the Lancet journals
- Associate Editor for BMC Genomics
- Grant reviewer: e.g. MRC, Wellcome Trust, Lundbeck Foundation
- PhD Examiner (> 10 PhDs to date)