

Shing Wan (Sam) Choi, PhD

Statistical Geneticist · Translational and Clinical Genomics · Polygenic Score Methodology

✉ choishingwan@gmail.com  github.com/choishingwan  gitlab.com/choishingwan ☎ (646)-797-0104

Summary

Creator of PRSice-2 and PRSet (2,000+ citations) and first author of the field-standard Nature Protocols PRS guide (2,500+ citations); second author on BridgePRS (Nature Genetics). At Regeneron: leads pharmacogenomic and proteomic analyses across 4 therapeutic areas and 10 clinical trials, generating genetic and proteomic evidence for clinical development decisions and building scalable, reproducible pipelines across the portfolio.

Professional Experience

Pharmacogenomics & Proteomics – Regeneron Genetics Center

Senior Manager (2024-Present); Manager (2022-2023)

- Lead the genetic and proteomic analysis strategy across 10 Phase 2/3 programs (ophthalmology, cardiovascular, metabolic, immunology), spanning active trial support and post-hoc analysis.
- Build and deploy end-to-end, cloud-scale polygenic score pipeline supporting multi-cohort training and validation, enabling reproducible genetics analyses across programs.
- Develop statistical framework for interpreting polygenic score results in underpowered clinical trials, triangulating across multiple polygenic score methods and leave-one-out sensitivity analyses to strengthen confidence in biological signal.
- Establish Olink proteomics analysis pipeline: plate-randomization, bridging-sample design, and normalization strategies to correct for batch and cross-cohort technical variation.
- Draft statistical analysis plans and partner with clinical and senior leadership to align on approach, ensuring on-time delivery against program milestones.

O'Reilly Lab – Icahn School of Medicine at Mount Sinai

Instructor (2021-2022); Postdoctoral Researcher (2019-2021)

- Developed PRSet, a proof-of-concept gene-set-based polygenic score method to segregate genetic signals into pathways and biological functions.
- Developed EraSOR, software that corrects for bias in polygenic score analyses introduced by sample overlap.

SGDP – King's College London

Postdoctoral Research Fellow (2016-2019)

- Developed PRSice-2, a widely adopted standalone software implementation of standard polygenic score methodology for biobank-scale data (cited 1.9k times as of Sept 2026).
- First author of the polygenic score tutorial, now a field-standard guideline (cited 2.5k times as of Sept 2026).

- Co-Director, Summer Workshop on Polygenic Risk Score

Technical Skills

- Domains: pharmacogenomics, polygenic risk prediction, clinical trial genetic analysis, statistical genetics, GWAS, multi-omics integration, high-throughput proteomics, pipeline engineering, reproducibility
- Languages and tools: R, Rust, Nextflow, WDL, C++, Python; Linux / HPC / AWS cloud

Education

The University of Hong Kong (2012-2016)

PhD in Bioinformatics

Thesis: “*The Genetics of Schizophrenia – Contribution of Common Variants and Genes Implicated in a Maternal Immune Activation Model.*” Supervisor: Prof. Pak Sham

The University of Hong Kong (2008-2011)

Bachelor of Science in Bioinformatics

Selected Publications

Full list of peer-reviewed publications available on Google Scholar

- **Choi SW**, Mak TS-H, O'Reilly PF. *Tutorial: a guide to performing polygenic risk score analyses*. Nature Protocols. (2020)
- **Choi SW**, O'Reilly PF. *PRSice-2: Polygenic Risk Score software for biobank-scale data*. GigaScience. (2019)
- Souaiaia T, Wu HM, Ori APS, **Choi SW**, Hoggart CJ, O'Reilly PF. *Striking Departures from Common-Variant Architecture in the Tails of Complex Traits*. Nature. (2026)
- Hoggart CJ, **Choi SW**, García-González J et al. *BridgePRS leverages shared genetic effects across ancestries to increase polygenic risk score portability*. Nat Genet (2024)
- **Choi SW**, Mak TSH, Hoggart CJ, O'Reilly PF. *EraSOR: a software tool to eliminate inflation caused by sample overlap in polygenic score analyses*. GigaScience. (2023)
- Mak TSH, Porsch RM, **Choi SW**, Zhou X, Sham PC. *Polygenic scores via penalized regression on summary statistics*. Genet Epidemiol. (2017)

Teaching & Mentorship

- Lecturer, Genetic Counseling, Icahn School of Medicine at Mount Sinai (2021)
- Lecturer, Statistical Genetics (MSc) and Genetic Epidemiology (Medical School), King's College London (2016-2019)
- Supervised 5 MSc and 1 PhD trainees (King's College London)
- Led NGS workshop for faculty at HKU Centre of Genomics Science