

# Dr Reju Sam John

Flat 2, 21 Aliford Avenue, One Tree Hill, 1061 — Auckland

(+64) 022 134 3292 | [rejusamjohn@gmail.com](mailto:rejusamjohn@gmail.com) | [LinkedIn: dr-reju-sam-john-14b00774](https://www.linkedin.com/in/dr-reju-sam-john-14b00774) | [ORCID: 0000-0002-5024-3700](https://orcid.org/0000-0002-5024-3700)

Permanent Resident of New Zealand | [Google Scholar: 169+ citations](#)

## Professional Summary

---

Computational epidemiologist and AI researcher. PhD-qualified, GCP R3 certified, with 7+ years delivering health data science and research coordination in academic and health settings.

Experienced in the full research lifecycle: study design, ethics applications, REDCap database management, data collection and analysis, and dissemination — across multidisciplinary teams of clinicians, researchers, and policy stakeholders.

Strong publication record in high-impact journals (*Journal of the Royal Society Interface*; *Nature Communications*; 169+ citations). Committed to rigorous, reproducible research practice and to the practical application of Te Tiriti o Waitangi principles in Aotearoa's health system. Keen interest in how AI can improve diagnostic services and health equity.

## Core Research Competencies

---

- AI / ML methods in health systems
- Research design & ethics applications
- Quantitative data collection & analysis
- REDCap database management
- Literature reviews & evidence synthesis
- Longitudinal & observational study coordination
- Multi-site, multidisciplinary project management
- Statistical modelling (Python, R)
- Health equity & Te Tiriti o Waitangi frameworks
- Stakeholder engagement & relationship management
- Scientific writing, reporting & dissemination
- ICH GCP R3 certification

## Relevant Work Experience

---

### AI, Machine Learning & Data Science Consultant

*Molecular Epidemiology and Public Health Laboratory*

Mar 2025 – Present

Remote (Palmerston North)

- Design and implement AI and ML pipelines (Python, R) to model infection dynamics, evaluate transmission risk, and produce evidence for public health recommendations.
- Translate analytical outputs into reports and presentations for clinical, policy, and non-technical audiences; mentor PhD researchers in reproducible research workflows.
- Apply Te Tiriti o Waitangi principles in practice: ensure Māori data sovereignty considerations are embedded in study design, data-sharing agreements, and dissemination planning; completed *Understanding Te Tiriti* (Groundwork, Sep 2025) to strengthen culturally grounded research practice.

### Postdoctoral Fellow — Data Science, Network Theory & Health Informatics

*The University of Auckland*

Jan 2023 – Mar 2024

Auckland

- Built and managed a longitudinal participant database for a clinical study tracking infection dynamics in vulnerable populations, integrating clinical records, laboratory results, and behavioural data while maintaining full audit-ready source documentation.
- Designed large-scale computational models of disease propagation coupled with (mis)information diffusion — pragmatic research designs directly relevant to AI-enabled health interventions and evidence generation.
- Developed reproducible, version-controlled analytical workflows in Linux environments and delivered results to academic, clinical, and policy audiences at national forums.

## Postdoctoral Fellow — Computational Epidemiologist & Health Data Scientist

Massey University

Nov 2020 – Jan 2023

Palmerston North

- Led the design and delivery of metapopulation SEIR models across 340 cities, producing the first-author paper published in *Journal of the Royal Society Interface* (2024) — demonstrating full end-to-end research ownership from conception to dissemination.
- Built scalable ETL pipelines integrating multi-source public health datasets (APIs, web scraping, administrative records), achieving 90% accuracy in epidemic trend forecasting using ensemble machine learning models.
- Developed and maintained structured epidemiological databases with quality control, data validation, and downstream AI-readiness, mirroring the data governance requirements of health AI evaluation research.
- Collaborated with international multidisciplinary teams (epidemiologists, virologists, clinicians, biostatisticians) across New Zealand, USA, and Europe.
- Contributed to ethics submissions, research reporting, and stakeholder briefings for Health Research Council-aligned projects.

## Postdoctoral Fellow — Astrophysics Simulations & Computational Data Science

Inter-University Centre for Astronomy and Astrophysics (IUCAA), India

Aug 2018 – Nov 2020

Pune, India

- Designed automated processing pipelines for large-scale (1 TB+) simulation datasets on HPC clusters; applied statistical inference and ML to high-dimensional data — foundational skills carried directly into subsequent health data science roles.

## Production Operator — Regulated Healthcare Manufacturing

Fisher & Paykel Healthcare

Sep 2024 – Feb 2026

East Tāmaki, Auckland

- Operated within an ISO-regulated medical device manufacturing environment, gaining direct operational knowledge of quality systems, risk documentation, and compliance requirements applicable to health technology evaluation.

## Selected Research & Analytics Projects

---

### AI Diagnostic Evaluation: Equity-Focused Framework for NZ Retinal Screening [github.com/reju...](https://github.com/reju...)

Simulation-based evaluation of AI diabetic retinopathy screening tools (incl. THEIA) with NZ equity analysis — directly targeting the AI Lab's research mandate.

- Simulated 10,000-patient NZ cohorts (ethnicity, NZDep quintile, age, sex) from published NZ prevalence data and evaluated three AI tool profiles under equal- and differential-performance scenarios for Māori, Pacific, and NZ European subgroups.
- Produced equity-stratified diagnostic accuracy metrics (sensitivity, specificity, AUC-ROC, calibration) with statistical testing for between-group differences. This methodology maps directly to how the AI Lab will evaluate live radiology and retinal screening AI tools for safe, equitable deployment.
- Structured as a reproducible modular pipeline (`config/`, `src/`, `notebooks/`) with all parameters cited from peer-reviewed literature.

### Healthcare Quality & Safety Insights Dashboard

[github.com/rejusam/...](https://github.com/rejusam/...)

Multi-page BI platform for health system adverse event, complaint, and KPI analytics — modelled on Health NZ Waitemata quality systems.

- Delivered seven interactive modules covering adverse events, complaints, KPI tracking, health equity metrics, SPC (3-sigma) charts, and automated stakeholder reporting.
- Implemented non-parametric hypothesis testing (Mann-Whitney U, Kruskal-Wallis, Chi-Square) and time series decomposition for evidence-based health performance analysis.

## Respiratory Biometrics Analytics Platform

[github.com/rejusam/ClinicalDataAnalysis](https://github.com/rejusam/ClinicalDataAnalysis)

End-to-end clinical analytics pipeline — 34,000+ ICU vital-sign measurements with full audit infrastructure and CI/CD.

- Four-gate ETL pipeline with audit logging, Power BI star-schema export, and interactive Streamlit dashboard; CI/CD via GitHub Actions with pytest data validation.

## Published Research Code: Infectious Disease Modelling

[SEIR\\_metapopulation](#) | [Modelling\\_Lassa](#)

Open-source repositories accompanying two first-author publications in *Journal of the Royal Society Interface* (2024). Metapopulation SEIR across 340 Chinese cities and zoonotic Lassa transmission modelling — demonstrating end-to-end reproducible research practice at publication standard.

## Education

---

**Ph.D. in Physics** (Computational Modelling, Data Analysis & Simulation)

Sep 2011 – Aug 2018

*Pondicherry University, India*

**M.Sc. in Physics**

Jun 2006 – Aug 2008

*Mahatma Gandhi University, India*

## Certifications

---

- **Understanding Te Tiriti o Waitangi (Concise)** — Groundwork (Sep 2025)
- **ICH Good Clinical Practice R3** — Global Health Training Centre (July 2025)
- **AWS Machine Learning Essential Training** (Mar 2025); **Google Cloud: Building Data Pipelines** (Apr 2025)
- **Deep Learning: Getting Started; NLP with Python for Machine Learning; Applied ML: Ensemble Learning** (Apr 2025)
- **Power BI: Dashboards for Beginners** (May 2024); **Hands-On Data Science using SQL, Tableau, Python, and Spark** (Mar 2024)

## Selected Publications (11 peer-reviewed articles; 169+ citations)

---

- **John, R.S.** et al. (2024). High connectivity and human movement limits the impact of travel time on infectious disease transmission. *J. Royal Society Interface*, 21(210). [doi:10.1098/rsif.2023.0425](https://doi.org/10.1098/rsif.2023.0425)
- **John, R.S.**, Fatoyinbo, H.O., & Hayman, D.T.S. (2024). Modelling Lassa virus dynamics in West African *Mastomys natalensis*. *J. Royal Society Interface*, 21(216). [doi:10.1098/rsif.2024.0106](https://doi.org/10.1098/rsif.2024.0106)
- Muylaert, R.L. et al. (2023). Identifying SARS-like coronavirus spillover risk hotspots. *Nature Communications*. [doi:10.1038/s41467-023-39567-2](https://doi.org/10.1038/s41467-023-39567-2)
- Hayman, D.T.S., **John, R.S.**, & Rohani, P. (2022). Transmission models indicate Ebola virus persistence in non-human primate populations is unlikely. *J. Royal Society Interface*, 19(187). [doi:10.1098/rsif.2021.0638](https://doi.org/10.1098/rsif.2021.0638)

Full list: [ORCID 0000-0002-5024-3700](https://orcid.org/0000-0002-5024-3700) | [Google Scholar](#)

## Additional Information

---

**Residency** Permanent Resident of New Zealand

**Languages** English (Fluent), Malayalam (Fluent), Te Reo Māori (Basic)

**Te Tiriti** Completed *Understanding Te Tiriti o Waitangi* (Groundwork, Sep 2025). Māori data sovereignty and equity principles embedded in current research design and dissemination practice.

**Referees** Available upon request

*Last updated: March 2026*