Dr Arif Budiarto

Postdoctoral Research Fellow, Usher Institute, University of Edinburgh
Email: arif.budiarto@ed.ac.uk | Phone: +44 744 7900 766
LinkedIn

RESEARCH PROFILE

Early-career researcher in medical informatics and health data science with specialised expertise in analysing large, real-world healthcare datasets at a national scale. My work spans electronic health records (EHR), genomics, and wearable data, with extensive hands-on experience accessing, processing, and modelling complex clinical datasets that are rarely available to early-career researchers in the UK. I focus on developing reproducible, clinically actionable prediction models for respiratory and chronic disease management in collaboration with NHS partners.

Key Research Strengths

- Extensive experience working with **large-scale, real-world NHS datasets, including OPCRD (**UK-wide primary care data), and **DataLoch** (Lothian health board).
- Proven ability to build and manage **end-to-end data pipelines for high-volume, high-dimensional** clinical data.
- Leadership in **multi-institutional data collaborations**, including GECCO (international genomics consortium) and the UK respiratory research project.
- Expertise in **methodological development** for prediction modelling using heterogeneous real-world data.

Key Research Metrics

- 70 peer-reviewed publications
- Over 1,000 citations
- Google Scholar h-index: 21
- Scopus h-index: 18

EDUCATION

PhD in Medical Informatics - The University of Edinburgh
December 2021 – January 2026

MSc in Advanced Design Informatics - The University of Edinburgh
September 2014 – June 2016
Award: Distinction

BSc in Computer Science - Bina Nusantara University
September 2007 - February 2011
GPA: 3.68 out of 4.00

CURRENT POSITION

Postdoctoral Research Fellow, University of Edinburgh (April 2026 – present)
I am applying data science and machine learning methods to large-scale linked health and social care data, with a focus on improving risk stratification and care pathways for individuals living with COPD. I am contributing to methodological development, analysis, and dissemination within a multidisciplinary programme aimed at advancing integrated, data-informed approaches to respiratory care.

RESEARCH EXPERIENCE AND LEADERSHIP

Postdoctoral Research Fellow, University of Edinburgh (August 2025 – March 2026)
Developed and validated risk prediction models for COPD outcomes using large-scale NHS datasets (primary and secondary care). Led data management, model development, and retrospective validation using collected clinical data. Working closely with NHS Lothian and funded by AstraZeneca.

PhD Researcher, University of Edinburgh (December 2021– January 2025)

- Developed machine-learning models for asthma attack prediction using OPCRD and DataLoch datasets.
- Led end-to-end data pipelines: extraction, cleaning, feature engineering, and supervised modelling.
- Produced first-author publications in JMIR AI, IEEE EMBC, and Pervasive Health.

Research Associate, Bioinformatics and Data Science Research Centre, Bina Nusantara University (October 2016 – December 2024)

- Managed data pipelines for colorectal cancer research, including GWAS modelling.
- Directed several machine-learning projects involving wearable data, telco, and agriculture data.

- Led genomics and clinical data analysis of Indonesian cohort for the Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO), contributing to multiple joint publications.

COLLABORATIONS AND INSTITUTIONAL INTEGRATION

- Active contributor to AUKCAR (now CARRii) and the Usher Institute.
- Close collaboration with NHS Lothian clinicians and data teams.
- Experience working with international partners, including GECCO and industry collaborators.

TECHNICAL SKILLS

- Python, R, SQL, C/C++, PHP, JavaScript
- Machine learning, EHR analytics, genomics pipelines, reproducible workflows
- API development, database management (MySQL, PostgreSQL)
- Project management (Jira, Trello), Confluence

SELECTED TRAINING

- Research, GDPR and confidentiality (2024)
- Generative AI for researchers (2024)
- Health Insurance Portability and Accountability Act of 1996 (HIPAA) (2020)
- Genomics Data Analysis (2020)

PUBLICATIONS

Publications
The full list is available on Scopus, Google Scholar, and ResearchGate.
Selected first-author papers include:

- JMIR AI (PhD work)
- IEEE EMBC 2023
- Pervasive Health 2024