

Linus Chirchir

Flat 13, 14 Heron Lane, Edinburgh, EH5 1FA

Mobile: 07522 113 567 E-mail: linus.chirchir@researchdata.scot

[ORCID](#) | [Google Scholar](#) | [Scopus](#) | [Web of Science](#)

Profile

Population and Health Data Scientist with experience in the preparation, linkage, and quantitative analysis of large-scale administrative and longitudinal datasets within health and social care research. Research interests include population health inequalities, life-course analysis, social determinants of health, and the application of advanced statistical methods to population-level data within Trusted Research Environments (TREs).

Employment

Data Analyst, [Research Data Scotland](#), Edinburgh, United Kingdom **09/2024 - Present**

- Establishing standards for data cleaning, quality assurance, metadata, and documentation to enable secure and FAIR access for demographic and social science research.
- Process sensitive datasets for the Scottish National Safe Haven as part of UK Research and Innovation (UKRI) funded Administrative Data Research programme.
- Develop and evaluate synthetic datasets using statistical modelling approaches (e.g., regression-based and tree-based methods), assessing fidelity, utility, and disclosure risk.

Health Data Analyst, [University of Edinburgh](#), Edinburgh, United Kingdom **09/2023 - 06/2024**

- Designed reproducible R and SQL data pipelines to integrate over 13,000 linked NHS orthopaedic records using regression modelling (linear and logistic) and record linkage methods to examine treatment pathways and patient outcomes.
- Developed interactive Shiny dashboard for orthopaedic outcomes to support clinical reporting and evidence-based decision making.
- Coordinated data governance processes, metadata standardisation and data dictionaries with the Data Controller and Research Database Manager for audit purposes and reuse of data.

ICT Officer/Researcher, [Moi Teaching & Referral Hospital](#), Eldoret, Kenya **07/2008 - 12/2021**

- Led data quality improvement and analytics initiatives across 104 administrative health datasets.
- Applied descriptive, correlation, regression, time-series analysis, and hypothesis testing to examine patterns in patient flow, service utilisation, workforce performance, trends in service activity, patient outcomes, and workforce metrics.
- Developed key performance indicators (KPIs) with clinical and administrative workforce to support healthcare delivery, resource allocation, and performance monitoring.
- Contributed to empirical research resulting in nine peer-reviewed publications.

Education

MSc in Health Data Science, [Swansea University](#), United Kingdom **01/2022 - 02/2023**

Dissertation: Evaluating Cognitive Assessments in the Progression to Early-Stage Alzheimer's Disease Using Machine Learning

- Developed and validated six machine learning models in R (Logistic Regression, KNN, SVM, Random Forest, GBM, AdaBoost) using the OASIS-3 dataset to predict early-stage Alzheimer's disease, achieving 91% accuracy and AUROC of 0.94.

Master of Business Administration (Health Leadership and Management), [United States International University-Africa](#), Kenya **09/2018 - 08/2019**

Project: Factors Influencing Electronic Health Record (EHR) Adoption by Nurses at Moi Teaching and Referral Hospital

- Conducted a quantitative cross-sectional study using survey data analysed in SPSS, applying correlation and linear regression techniques to assess determinants of EHR adoption to provide recommendations to support digital health implementation, including workforce training.

Master of Information Systems, [Kisii University](#), Kenya **01/2014 - 12/2015**

Thesis: Effects of Health Information Systems on User Performance: A Survey of Health Organisations in Uasin Gishu County, Kenya

- Applied descriptive statistics, correlation analysis, multivariable regression modelling, and hypothesis testing to examine direct and mediated relationships between system characteristics and user performance.

Bachelor of Business Management, [Moi University](#), Kenya **09/2008 - 12/2012**

- Developed foundational training in statistics, database systems, programming, and operations research, supporting later work in data management, statistical analysis, and health data science.

Grant Funding

Principal Investigator - £64,508.43

- Awarded funding for *FRAME-DS: Framework for Recognising Applied Data Stewardship Within TRE Workforces – A Collaborative, Cross-Organisation Pilot* through the UKRI Careers and Skills for Data-Driven Research (CaSDaR) Network+ Research Technical Professional Call (part of Grant No. UKRI739), March 2026.

Honours and Recognition

UK Research and Innovation (UKRI) Global Talent Endorsement (May 2026)

- Endorsed by UK Research and Innovation under the Global Talent (Endorsed Funder) route in recognition of contribution to UKRI-funded research and innovation activities.

Technical Skills

- **Programming and Statistical Analysis:** R, SQL, SPSS, Advanced Excel
- **Statistical Methods:** Longitudinal and cohort analysis, demographic and population data analysis, regression modelling (linear, logistic, Poisson, negative binomial, Cox proportional hazards, multilevel/mixed effects), epidemiological analysis, causal inference, record linkage, machine learning, spatial analysis, and hypothesis testing.
- **Data Visualisation and GIS:** R, Tableau, Power BI, QGIS, ArcGIS
- **Reproducible Research and Version Control:** Quarto, Git, GitHub

Peer Review and Research Engagement

- **ESRC Peer Review College**
Expert reviewer for population methods, data science and social statistics proposals.
- **Peer Reviewer**
Reviewed 19 manuscripts across 7 journals, including the International Journal of Population Data Science, Journal of Medical Internet Research (JMIR), JMIR Medical Informatics, JMIR Bioinformatics and Biotechnology, JMIR Perioperative Medicine and PLOS ONE.
- **Mentor, The Mentoring Club**
Provide one-to-one mentoring to early-career professionals in data science and analytics, offering guidance on research methods, career development and professional practice.
- **Publons Academy Supervisor (Peer Review Mentor)**
Mentored 6 trainees on the Publons Practical Peer Review Training Course, providing structured feedback and final review assessments.

Selected Peer Reviewed Conference Presentations and Journal Articles

- [Automating Low-Fidelity Synthetic Data Generation from Metadata: A Reproducible Approach Using R](#). Proceedings of the Administrative Data Research UK Conference 2025.
- [ADRUK Ingest Improvement Task and Finish Group: Smoothing the ingest of data](#). Proceedings of the Administrative Data Research UK Conference 2025.
- [Change Process Factors Influencing Electronic Health Records Adoption by Nurses at Moi Teaching and Referral Hospital, Kenya](#). HCist - International Conference on Health and Social Care Information Systems and Technologies 2020.
- [Determining Role of Social Factors in Adoption of Electronic Health Records by Nurses](#). International Conference on Informatics, Management, and Technology in Healthcare 2020.
- **Chirchir, L. K.,** Aruasa, W. K., & Chebon, S. K. (2021). [Technical Factors Influencing Electronic Health Records Adoption by Nurses at Moi Teaching and Referral Hospital, Kenya](#). African Journal of Science, Technology, Innovation and Development, 13(5), 551 - 558.
- Aruasa, W. K., **Chirchir, L. K.,** & Chebon, S. K. (2019). [Determinants of Physicians and Nurses' Professional Satisfaction at the Moi Teaching and Referral Hospital, Eldoret, Kenya](#). *Journal of Health, Medicine and Nursing*, 64, 6 – 17.

Working Papers / Current Research

- Evaluating Cognitive Assessments in Early-Stage Alzheimer's Disease Using Machine Learning. Manuscript in Preparation.
- Synthetic Data Quality Assessment for Administrative and Population Health Data.
- FRAME-DS: Applied Data Stewardship in Trusted Research Environments.
- Metadata-Driven Synthetic Data Generation for Administrative Data Research.

Certifications and Training

- EPSRC and ESRC Peer Review Training, November 2025
- Office for National Statistics Safe Researcher Training (ONS-SRT), June 2025
- Certified Publons Academy Peer Reviewer (September 2020) and Mentor (April 2021)

Memberships and Fellowships

- Research Data Alliance (RDA) (Member)
- The Federation for Informatics Professionals (Advanced Practitioner)
- Royal Statistical Society (Member and Graduate Statistician)
- Association of Professional Healthcare Analysts (Member)