

# Linxuan Zhang

linxuan.z@outlook.com | linxuanzhang.com | Google Scholar | LinkedIn

## Profile

---

Researcher and data scientist with seven years of industry experience and a quantitative background spanning mathematics, statistics, machine learning and causal inference. Currently pursuing a PhD in Biostatistics at the University of Cambridge while serving as Artificial Intelligence Lead at Sequoia Genetics. Research focuses on integrating large-scale genetic and molecular data, with particular emphasis on AI, statistical genetics and causal methods for drug-target discovery and validation.

## Experience

---

**Artificial Intelligence Lead** **Oct. 2024 – Present**  
*Sequoia Genetics* *London, UK*

- Design and implement statistical, causal-inference and machine-learning methodologies using human genetic data to inform drug-target discovery and development for pharmaceutical, biotechnology, investment and academic clients.
- Lead development of reproducible AI and analytical workflows that translate methodological research into practical evidence for target identification and validation.
- Contributed applied analyses to a collaboration with the University of Cambridge and Eli Lilly evaluating methods for index-event bias in Mendelian-randomisation studies of disease progression.

**Data Scientist** **Sep. 2019 – Sep. 2024**  
*Lane Clark & Peacock, Health Analytics* *London, UK*

- Delivered statistical and machine-learning analyses across health, insurance and energy, including time-series forecasting for wind generation and trend analysis for reserving.
- Built production data pipelines and ETL processes, and analysed large, complex datasets to turn research questions into repeatable analytical products.
- Led an LLM-based project to construct a previously unavailable medical database; also delivered internal training in Python, R, statistics, machine learning and version-control practices.

## Education

---

**PhD in Biostatistics** **Oct. 2025 – Expected 2028**  
*University of Cambridge, MRC Biostatistics Unit* *Cambridge, UK*

- Supervisors: Professor Stephen Burgess and Dr Dipender Gill. Industrial CASE studentship with Sequoia Genetics.
- Research: statistical and machine-learning methods for integrating genetic and molecular data, combining AI and causal inference to inform drug discovery.

**MSc, Data Science and Machine Learning (part-time)** **2021 – 2023**  
*Imperial College London* *London, UK*

**MMath, Mathematics and Statistics** **2015 – 2019**  
*University of Oxford, Keble College* *Oxford, UK*

## Selected Publications

---

**Zhang, L.<sup>†</sup>, Higgins, I. A.<sup>†</sup>, Dai, Q., et al. (2026).** “Comparison of methods for assessing effects of risk factors on disease progression in Mendelian randomization under index event bias.” *The American Journal of Human Genetics*, 113. First author and equal contributor; led applied analyses. doi:10.1016/j.ajhg.2026.06.004

**Huang, J., Zhang, L., Bodimeade, C., et al. (2025).** “Unravelling the Relationship Between Height, Lean Mass, Alzheimer’s Disease and Cognition Through Mendelian Randomization.” *Genes*, 16(2), 113. doi:10.3390/genes16020113

<sup>†</sup>Equal contribution. Full publication list available on Google Scholar.

## Earlier Research Experience

---

|                                                                                               |                                                        |
|-----------------------------------------------------------------------------------------------|--------------------------------------------------------|
| <b>Artificial Intelligence Research Assistant</b><br><i>University of Southern California</i> | <b>Aug. 2018 – Oct. 2018</b><br><i>Los Angeles, US</i> |
|-----------------------------------------------------------------------------------------------|--------------------------------------------------------|

- Developed behavioural-sociology models and tree-search algorithms.

|                                                                                                          |                                                   |
|----------------------------------------------------------------------------------------------------------|---------------------------------------------------|
| <b>Artificial Intelligence Research Assistant</b><br><i>University of Oxford, Mathematical Institute</i> | <b>Jun. 2018 – Aug. 2018</b><br><i>Oxford, UK</i> |
|----------------------------------------------------------------------------------------------------------|---------------------------------------------------|

- Applied transfer learning with VGG-16 to an image-classification research project.

## Technical Expertise

---

|                                |                                                                                                                                                      |
|--------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Statistical methods</b>     | Causal inference, Mendelian randomisation, statistical genetics, uncertainty quantification, simulation, experimental and observational study design |
| <b>AI and machine learning</b> | Supervised and unsupervised learning, deep learning, LLM and agentic workflows, model evaluation, feature engineering                                |
| <b>Programming and data</b>    | Python, R, SQL, PySpark, pandas, NumPy, scikit-learn, TensorFlow, reproducible pipelines, ETL, version control                                       |
| <b>Communication</b>           | Technical writing, client-facing analysis, research translation, training and interdisciplinary collaboration                                        |

## Qualifications and Interests

---

|                           |                                                                                                                                 |
|---------------------------|---------------------------------------------------------------------------------------------------------------------------------|
| <b>Professional</b>       | CFA Level I passed; TensorFlow Developer Certificate                                                                            |
| <b>Research interests</b> | AI for biology, statistical genetics, drug-target validation, prediction versus intervention, decision-making under uncertainty |
| <b>Broader interests</b>  | Value investing, accounting, business analysis, music composition and performance, strategy games                               |