**FEIQING HUANG (Amiee)**
Translational data scientist applying LLMs, causal inference, and longitudinal EHR to real-world evidence and clinical decision support.

655 Huntington Ave • Boston, MA 02115 • fqhuang@hsph.harvard.edu • 857-445-8803

## Education

**Harvard T.H. Chan School of Public Health** — Boston, MA
Postdoc, Biostatistics — August 2023 – Present
Advised by Prof. Tianxi Cai.

**Harvard-MIT Center for Regulatory Science** — Boston, MA
Regulatory Science Fellow — October 2024 – Present
Advised by Dr. Florence Bourgeois.

**University of Hong Kong** — Hong Kong SAR
PhD, Statistics — September 2018 – May 2023
Thesis: *Modeling sequential dependence in statistics and machine learning.*
Advised by Prof. Guodong Li. Recipient, Hong Kong PhD Fellowship (2018).

**University of Hong Kong** — Hong Kong SAR
B.Sc. in Statistics and Quantitative Finance — September 2014 – May 2018
GPA: 3.97 (Top 5 in major). Scholarships: Dr Patrick S C Poon, Saw Swee Hock, C.V. Starr, Statistics & Actuarial Science, Dean's Honor List.

**UC Davis** — Davis, CA
Coursework: Statistics — September 2016 – December 2016

## Research Experience

**Harvard T.H. Chan School of Public Health, Dept. of Biostatistics** — Boston, MA
*Postdoctoral Researcher* — Aug 2023 – Present

- Independently design and execute translational ML studies using longitudinal EHR, clinical notes, and biomedical knowledge graphs; deliver interpreted insights to clinicians and data scientists.
- Led a NeurIPS 2024 project integrating LLMs with biomedical knowledge graphs via resource-efficient co-fine-tuning to improve cross-modal clinical representation learning; presented results to cross-functional audiences.
- Building a privacy-preserving synthetic longitudinal clinical notes pipeline (diffusion models + open-source LLMs) conditioned on clinical knowledge graphs to enable training foundation clinical LLMs and benchmarking agent-based reasoning systems.
- Developing reinforcement-learning strategies for budget-constrained sequential feature acquisition to reduce manual chart review in phenotyping workflows.

**Harvard–MIT Center for Regulatory Science** — Boston, MA
*Regulatory Science Fellow* — Oct 2024 – Present

- Execute U01-funded real-world evidence studies with FDA, Mass General Brigham, and UPMC using longitudinal EHR to evaluate comparative effectiveness in MS and RA.
- Authored a reproducible R/Python pipeline for cohort construction, outcome imputation, and causal analyses with calibration; tutorial published in *JMIR* (2025); code on GitHub.
- Communicate complex findings clearly to regulators, and clinicians; tailor deliverables to decision-making needs.

**The University of Hong Kong** — Hong Kong SAR
*Ph.D. Researcher, Statistics* — Sep 2018 – May 2023

- Designed a recurrent self-attention architecture for long-sequence modeling (ICLR 2023, top 5%); applications across forecasting, code analysis, and NLP.
- Developed high-dimensional time-series models and tensor methods; work published in *Journal of Econometrics* and other venues.

- Mentored junior researchers; coordinated paper submissions and presentations.

## Publications

**Deep Learning Applications**

**Huang, F.**, Zhang, S., Sweet, S. & Cai, T. (2024). A Teacher-teacher Framework for Clinical Language Representation Learning. NeurIPS 2024.

**Huang, F.**, Lu, K., Cai, Y., Qin, Z., Fang, Y., Tian, G., & Li, G. (2023). Encoding Recurrence into Transformers. ICLR 2023.

Zhao, J., **Huang, F.**, Lv, J., Duan, Y., Qin, Z., Li, G., & Tian, G. (2020). Do RNN and LSTM have Long Memory? ICML 2020.

Wang, D., **Huang, F.**, Zhao, J. & Li, G. (2020). Compact Autoregressive Network. AAAI 2020.

**Huang, F.**, Si, Y., Zheng, Y., & Li, G. (2021). A New Measure of Model Redundancy for Compressed CNNs. Arxiv.

**Clinical Applications**

**Huang, F.**, Hou, J., Zhou, N., Greco, K., Lin, C., Sweet, S., Wen, J., Shen, L., Gonzalez, N., Zhang, S., Liao, K., Cai, T., Xia, Z., Bourgeois, F., & Cai, T. (2025). Advancing the Use of Longitudinal EHR: Tutorial for Uncovering RWE in Chronic Disease Outcome. *JMIR*, 27, e71873.

**Huang, F.**, Zhu, W., Hou, J., Sweet, S. M., Han, Y., Wen, J., Liao, K. P., Cai, T., Chitnis, T., Bourgeois, F., Xia, Z., & Cai, T. (2025). Comparing Ocrelizumab versus Natalizumab: A Real-World Evidence Study using EHR Disability Outcomes. [Manuscript].

Hou, J., **Huang, F.**, McDermott, G. C., Wen, J., Jeffway, M. I., Qi, Y., Han, Y., Cai, T., Liao, K. P., Bourgeois, F., & Cai, T. (2025). Real-world Study of Rheumatoid Arthritis Therapies Using Machine Learning Imputed Disease Activity Scores: A Retrospective Observational Study. [Manuscript].

**Statistical Theories**

Cai, T., **Huang, F.**, Nakada, R., Zhang, L. & Zhou, D. (2025). Contrastive Learning on Multimodal Analysis of EHR. *Journal of the American Statistical Association* (in revision).

**Huang, F.**, Rong, M. & Cai, T. (2025). Enhancing Spectral Embedding through Robust and Flexible Knowledge Transfer in Electronic Health Records. *Journal of the American Statistical Association* (submitted).

**Huang, F.**, Lu, K. & Li, G. (2025). High-Dimensional Low-Rank Linear Time Series Modeling. *Journal of Econometrics*, 249, 105995.

Cai, Y., Li, L., Lu, K., **Huang, F.**, Yao, Z. & Li, G. (2025). Unraveling Recurrent Dynamics. *Journal of the American Statistical Association AI special issue* (in revision).

**Huang, F.**, Zheng, Y., Lu, K. & Li, G. (2025). SARMA: Scalable High-Dimensional VARMA Model. *Statistical Sinica* (in revision).

## Leadership & Activities

**HKU Alumni Association of New England** — Boston, MA

Co-convener — December 2024 – Present

- Organized networking and social events for HKU alumni and visiting students in New England.
- Facilitated mentorship programs connecting exchange students with local professionals.

**Habitat for Humanity** — Hong Kong SAR

Event Organizer — 2015

- Co-organized spring cleaning events for the elderly; led volunteer recruitment and coordination.

## Skills & Interests

**Technical:** R (tidyverse, data.table, R Markdown), Python (pandas, NumPy, scikit-learn, PyTorch, HF), Jupyter.
**Languages:** English (Proficient), Mandarin (Native), Cantonese (Native), French (A2–B1)
**Interests:** Tennis, swimming, painting, reading